Automatica 93 (2018) 12–19

Contents lists available at ScienceDirect

Automatica

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

A necessary and sufficient condition for stability of LMS-based consensus adaptive filters ☆

Siyu Xie, Lei Guo *

Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
School of Mathematical Science, University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:
Received 4 January 2017
Received in revised form 10 September 2017
Accepted 28 December 2017

Keywords:
Least mean squares
Adaptive filters
Consensus strategies
Exponentially stable
Stochastic averaging
Switching networks
Performance analysis

ABSTRACT

This paper investigates the stability and performance of the standard least mean squares (LMS)-based consensus adaptive filters under a changing network topology. We first analyze the stability for possibly unbounded, non-independent and non-stationary signals, by introducing an information condition that can be shown to be not only sufficient but also necessary for the global stability. We also demonstrate that the distributed adaptive filters can estimate a dynamic process of interest from noisy measurements by a set of sensors working in a collaborative manner, in the natural scenario where none of the sensors can fulfill the estimation task individually. Furthermore, we give an analysis of the filtering error under various assumptions without stationarity and independency constraints on the system signals, and thus do not exclude applications to stochastic systems with feedback. In contrast to the analyses of the normalized LMS-based distributed adaptive filters, we need to use stochastic averaging theorems in the stability analysis due to possible unboundedness of the system signals.

1. Introduction

When filtering or tracking an unknown signal or parameter process in a distributed sensor network, each sensor can produce a local estimate based on its own noisy measurements and on the information gathered from other sensors. There are essentially three scenarios for this problem, i.e., centralized processing, distributed processing and the combination of both. In the first scenario, all the sensors transmit information to a fusion center, while in the second scenario, no fusion center is required and any sensor will conduct the estimation task through communicating with its neighbors. For centralized processing, collecting measurements from all other distributed sensors over the network may not be feasible in many practical situations due to limited communication capabilities, energy consumptions, packet losses or privacy considerations. These are the main motivations for the development of the distributed algorithms, in which any sensor only needs to exchange information with its neighbors, which will be more robust, need fewer communications and allow parallel information processing.

There are basically three types of strategies for distributed algorithms in the literature, namely, incremental strategies (Cattivelli & Sayed, 2011; Lopes & Sayed, 2007), consensus strategies (Braca, Marano, & Matta, 2008; Chen, Wen, Hua, & Sun, 2014; Kar & Moura, 2011; Solo, 2015) and diffusion strategies (Cattivelli & Sayed, 2010; Chen & Sayed, 2015a, b; Nosrati, Shamsi, Taheri, & Sedaaghi, 2015; Piggott & Solo, 2016, 2017; Sayed, 2014a, b). Despite of extensive researches on distributed adaptive filtering algorithms in recent years, most of the existing literature on stability and performance analyses require statistical independency or stationarity assumptions for the system signals (e.g., Cattivelli & Sayed, 2010, 2011; Kar & Moura, 2011; Lopes & Sayed, 2007; Nosrati et al., 2015; Sayed, 2014a, b), which cannot be satisfied in many practical situations, for example, signals from feedback systems. Some other interesting works that do not require the temporal independency of the regressors can be found when the signal to be estimated is a constant (Chen & Sayed, 2015a; Piggott & Solo, 2016, 2017; Solo, 2015). In particular, Piggott and Solo (2016) appear to be the first to study the almost sure convergence of the LMS-based distributed algorithms, and Piggott and Solo (2017) establish the second order performance results under temporally strictly stationary and strong mixing assumptions.

To provide stability and performance analyses under more general correlated and non-stationary situations, Chen, Liu, and

☆ This work was supported by the National Natural Science Foundation of China under grants 11688101 and 61227902. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Er-Wei Bai under the direction of Editor Torsten Söderström.

* Corresponding author at: Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.

E-mail addresses: xiesiyu@amss.ac.cn (S. Xie), lguo@amss.ac.cn (L. Guo).

https://doi.org/10.1016/j.automatica.2018.03.027

Guo (2014) considered a normalized diffusion LMS algorithm under a cooperative stochastic information condition with a fixed topology, but without independency or stationarity considerations. Their stability results and analyses were later improved (Chen, Liu, & Guo, 2016). In our recent work (Xie & Guo, 2015), we analyzed the stability of normalized consensus LMS algorithm with a fixed topology, under a more general information condition than that used in Chen et al. (2014, 2016), which has also been shown to be necessary for the stability of the consensus algorithm in a certain sense. However, the random matrix product methods (Chen et al., 2014, 2016; Xie & Guo, 2015) for the normalized LMS fail to be applicable to the standard LMS consensus adaptive filtering algorithms, because of the possible unboundedness of the system signals. This is one of the key issues that we have to deal with for the standard LMS filtering algorithms, and a preliminary step was recently made on stability analysis with a fixed network topology (Xie & Guo, 2017).

The main contributions of the paper contain the following three aspects: (i) We will present a weakest possible information condition for the stability of the standard LMS-based consensus adaptive filters under possibly unbounded, non-independent and non-stationary assumptions, which does not exclude applications to stochastic systems with feedback. Stochastic averaging theorems will be used in the stability analysis. (ii) We will show that the whole sensor network can accomplish the estimation task cooperatively, even if none of the sensors can do it individually due to lack of sufficient information, and we will also give a performance analysis for the mean square tracking error matrix under some mild assumptions. (iii) We allow the network topology to change over time and be jointly connected, which is applicable to situations where communication interruptions may happen between sensors.

In the rest of the paper, we will present the consensus adaptive filters based on the standard LMS in Section 2. Some necessary notations, concepts and mathematical definitions are stated in Section 3. The main results and proofs are given in Sections 4 and 5, respectively. Section 6 gives a simulation result and Section 7 concludes the paper and discusses related future problems.

2. Problem formulation

In this paper, we assume that the signal model at any sensor i $(i = 1, \dots, n)$ of the sensor network is described by a stochastic time-varying linear regression as follows

$$y_k^i = (\boldsymbol{\varphi}_k^i)^T \boldsymbol{\theta}_k + v_k^i, \quad k \geq 0, \tag{1}$$

where y_k^i is the scalar observation of the sensor i at time k, v_k^i is the disturbance or un-modeled dynamics, $\boldsymbol{\varphi}_k^i$ is the $m \times 1$-dimensional stochastic regressor of the sensor i, and $\boldsymbol{\theta}_k$ is an unknown $m \times 1$-dimensional stochastic signals whose variation at time k is denoted by $\boldsymbol{\omega}_k$, i.e.,

$$\boldsymbol{\omega}_k \triangleq \boldsymbol{\theta}_k - \boldsymbol{\theta}_{k-1}, \quad k \geq 1. \tag{2}$$

To estimate the unknown $\{\boldsymbol{\theta}_k, k \geq 0\}$, we consider the standard LMS-based consensus adaptive filter, which is recursively defined at each sensor $i = 1, \dots, n$ as follows

$$\begin{aligned} \widehat{\boldsymbol{\theta}}_{k+1}^i = \widehat{\boldsymbol{\theta}}_k^i + \mu \Bigg\{ & \boldsymbol{\varphi}_k^i [y_k^i - (\boldsymbol{\varphi}_k^i)^T \widehat{\boldsymbol{\theta}}_k^i] \\ & - \nu \sum_{l \in \mathcal{N}_{i,k}} a_{li,k} (\widehat{\boldsymbol{\theta}}_k^i - \widehat{\boldsymbol{\theta}}_k^l) \Bigg\}, \quad k \geq 0, \end{aligned} \tag{3}$$

where $\nu \in (0, 1)$ is a weighting constant, $\mu \in (0, 1)$ is the step-size, $\{a_{li,k}\}$ is the adjacency matrix of the sensor network, and $\mathcal{N}_{i,k}$ is the set of neighbors of the sensor i at time k (see the next section for details). We remark that if $\nu = 0$, then the above algorithm reduces to n independent LMS filters, which have been extensively studied in the literature (see e.g., Guo, Ljung, & Wang, 1997; Solo & Kong, 1995; Widrow & Stearns, 1985).

To write the above distributed adaptive filters into a compact form, we introduce the following notations:

$$\begin{aligned} &\boldsymbol{Y}_k \triangleq \mathrm{col}\{y_k^1, \dots, y_k^n\}, \quad \boldsymbol{\Phi}_k \triangleq \mathrm{diag}\{\boldsymbol{\varphi}_k^1, \dots, \boldsymbol{\varphi}_k^n\}, \\ &\boldsymbol{\Omega}_k \triangleq \mathrm{col}\{\underbrace{\boldsymbol{\omega}_k, \dots, \boldsymbol{\omega}_k}_{n}\}, \quad \boldsymbol{\Theta}_k \triangleq \mathrm{col}\{\underbrace{\boldsymbol{\theta}_k, \dots, \boldsymbol{\theta}_k}_{n}\}, \\ &\boldsymbol{V}_k \triangleq \mathrm{col}\{v_k^1, \dots, v_k^n\}, \quad \widehat{\boldsymbol{\Theta}}_k \triangleq \mathrm{col}\{\widehat{\boldsymbol{\theta}}_k^1, \dots, \widehat{\boldsymbol{\theta}}_k^n\}, \\ &\widetilde{\boldsymbol{\Theta}}_k \triangleq \mathrm{col}\{\widetilde{\boldsymbol{\theta}}_k^1, \dots, \widetilde{\boldsymbol{\theta}}_k^n\}, \ \text{where } \widetilde{\boldsymbol{\theta}}_k^i = \widehat{\boldsymbol{\theta}}_k^i - \boldsymbol{\theta}_k, \\ &\boldsymbol{F}_k \triangleq \mathrm{diag}\{\boldsymbol{\varphi}_k^1(\boldsymbol{\varphi}_k^1)^T, \dots, \boldsymbol{\varphi}_k^n(\boldsymbol{\varphi}_k^n)^T\}, \\ &\boldsymbol{G}_k \triangleq \boldsymbol{F}_k + \nu(\mathcal{L}_k \otimes I_m) \end{aligned}$$

where $\mathrm{col}\{\cdots\}$ denotes a vector by stacking the specified vectors, $\mathrm{diag}\{\cdots\}$ is used in a non-standard manner which means that $m \times 1$ column vectors are combined "in a diagonal manner" resulting in an $mn \times n$ matrix, $\mathcal{L}_k$ is the Laplacian matrix of the graph at time k (see the next section for details), $\otimes$ is the Kronecker product, and I_m denotes the m-dimensional identity matrix. By (1) and (2), we have

$$\boldsymbol{Y}_k = \boldsymbol{\Phi}_k^T \boldsymbol{\Theta}_k + \boldsymbol{V}_k, \quad k \geq 0,$$

and

$$\boldsymbol{\Omega}_{k+1} = \boldsymbol{\Theta}_{k+1} - \boldsymbol{\Theta}_k, \quad k \geq 0.$$

From (3), we obtain that $\forall k \geq 0$,

$$\widehat{\boldsymbol{\Theta}}_{k+1} = \widehat{\boldsymbol{\Theta}}_k + \mu \boldsymbol{\Phi}_k (\boldsymbol{Y}_k - \boldsymbol{\Phi}_k^T \widehat{\boldsymbol{\Theta}}_k) - \mu\nu(\mathcal{L}_k \otimes I_m)\widehat{\boldsymbol{\Theta}}_k,$$

where $\mathcal{L}_k$ is the Laplacian matrix of graph $\mathcal{G}_k$ (see the next section). Let us denote $\widetilde{\boldsymbol{\Theta}}_k = \widehat{\boldsymbol{\Theta}}_k - \boldsymbol{\Theta}_k$ and because $(\mathcal{L}_k \otimes I_m)\boldsymbol{\Theta}_k = 0$, we can get $\forall k \geq 0$,

$$\widetilde{\boldsymbol{\Theta}}_{k+1} = (I_{mn} - \mu \boldsymbol{G}_k)\widetilde{\boldsymbol{\Theta}}_k + \mu \boldsymbol{\Phi}_k \boldsymbol{V}_k - \boldsymbol{\Omega}_{k+1}. \tag{4}$$

Note that by the stochastic internal–external stability results (see Propositions 2.1 and 2.2 in Guo, 1994), we know that the stability of (4) essentially hinges on the exponential stability of the homogeneous part:

$$\widetilde{\boldsymbol{\Theta}}_{k+1} = (I_{mn} - \mu \boldsymbol{G}_k)\widetilde{\boldsymbol{\Theta}}_k. \tag{5}$$

This motivates us to give some definitions on exponential stability in the next section.

3. Notations and definitions

Notations. Let $\mathbb{R}^{m \times n}$ denote the set of $m \times n$ matrices with real entries. For any random matrix $X \in \mathbb{R}^{m \times n}$, its Euclidean norm is defined as its maximum singular value, i.e., $\|X\| = \{\lambda_{max}(XX^T)\}^{\frac{1}{2}}$, where $\lambda_{max}(\cdot)$ denotes the largest eigenvalue of matrix $(\cdot)$ and $(\cdot)^T$ denotes the transpose operator, and its L_p-norm is defined as $\|X\|_{L_p} = \{\mathbb{E}[\|X\|^p]\}^{\frac{1}{p}}$, where $\mathbb{E}[\cdot]$ denotes the expectation operator.

Network topology. Consider a set of n sensors and model it as an undirected weighted graph. Since the relationship between neighbors may change over time, so does the graph describing it. Then we have a class of graphs $\mathcal{G}_k$ on the n vertexes at time k $(k \geq 0)$ composed of $\{\mathcal{V}, \mathcal{E}_k, \mathcal{A}_k\}$, where $\mathcal{V} = \{1, 2, \dots, n\}$ is the vertex set, $\mathcal{E}_k \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set at time k, and $\mathcal{A}_k$ is a matrix reflects the interaction strength among neighboring vertexes at time k. Define $\mathcal{A}_k = \{a_{ij,k}\}_{n \times n}$, $i, j = 1, \dots, n$, $k \geq 0$, which is called the weighted adjacency matrix with $a_{ij,k} \geq 0$, $\sum_{j=1}^n a_{ij,k} = 1$, $\forall i = 1, \dots, n$, $k \geq 0$. Since the graph $\mathcal{G}_k$ is undirected, we have $a_{ij,k} = a_{ji,k}$. Vertex i denotes the ith sensor and (i, j) denotes the connection from sensor