Signal Processing 122 (2016) 93–97

Contents lists available at ScienceDirect

Signal Processing

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

Fast communication

Variable step-size diffusion least mean fourth algorithm for distributed estimation

Jingen Ni [a,*], Jian Yang [a,b]

[a] School of Electronic and Information Engineering, Soochow University, Suzhou 215006, China
[b] Seagate Technology (Suzhou) Co., Ltd., Suzhou 215021, China

ARTICLE INFO

Article history:
Received 27 August 2015
Received in revised form
8 November 2015
Accepted 24 November 2015
Available online 11 December 2015

Keywords:
Diffusion strategy
Distributed estimation
Variable step-size
Least mean fourth algorithm

ABSTRACT

The diffusion LMS (DLMS) is one of the most popular online distributed estimation algorithms, due to its simplicity and ease of implementation. However, it may suffer from large steady-state misalignment in some strong, non-Gaussian noise environments. To address this problem, this paper introduces a diffusion least mean fourth (DLMF) algorithm by using the mean-fourth error cost function in a diffusion strategy. Moreover, a variable step-size (VSS) method is developed to further reduce the steady-state misalignment of the DLMF. Simulation results show that the DLMF outperforms the DLMS with uniform or binary noise, and that the VSS-DLMF has a superior steady-state performance as compared to the DLMF.

1. Introduction

In centralized estimation, agents in the network transmit their data to a fusion center for processing, which needs extensive amounts of communications between the fusion center and agents. Besides, this kind of network is not robust against the failure of the fusion center. To address this problem, the concept of distributed estimation was proposed. Distributed estimation can be used in scenarios where a set of agents are required to collectively estimate some unknown vector without using a fusion center [1].

In recent years, distributed estimation has been widely used in many applications [2], and many distributed estimation algorithms were proposed, such as the incremental LMS [3,4], incremental APA [5,6], diffusion LMS (DLMS) [7,8], and multitask DLMS [9,10]. Moreover, some distributed stochastic gradient (sub-gradient) descent algorithms for solving a general convex cost function were developed, e.g., in [11,12]. Among these algorithms, the DLMS [8] is one of the most popular online distributed estimation algorithms, due to its simplicity and ease of implementation. Recently, two DLMS with reduced communication overhead were proposed in [13,14], respectively. Although the DLMS and its variants have many advantages, in the case where the measurement noise is non-Gaussian, such as strong uniform or binary noise, it may suffer from high steady-state misalignment.

Research has shown that adaptive algorithms based on high-order moment cost function may yield lower steady-state misalignment than those based on mean-square error (MSE) one in some strong, non-Gaussian noise environments. Among others, the well-known least mean fourth (LMF) is a typical adaptive algorithm derived form high-order moment cost function minimization [15]. Over the last decade, the performance and variants of the LMF have been extensively studied, e.g., [16–20].

This paper first introduces a diffusion LMF (DLMF) algorithm by employing the mean-fourth error (MFE) cost function in the diffusion strategy presented in [12] to enhance the performance of distributed estimation in some strong, non-Gaussian noise environments. Then a

* Corresponding author. Tel.: +86 18260117775.
E-mail address: jni@suda.edu.cn (J. Ni).

http://dx.doi.org/10.1016/j.sigpro.2015.11.020
0165-1684/

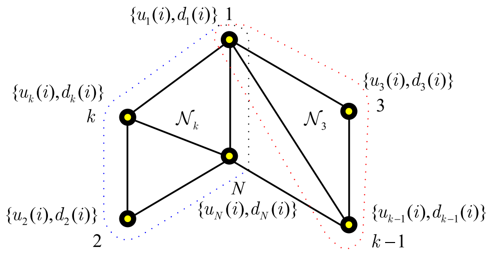

Fig. 1. A diffusion network consisting of N agents, where $u_k(i)$ and $d_k(i)$ represent the input and observation signals of agent k, respectively.

variable step-size (VSS) method is developed for the DLMF to address the tradeoff between fast convergence rate and low steady-state misalignment. Simulation results will show that when the measurement noise is strong uniform or binary noise, the DLMF obtains lower steady-state misalignment than the DLMS and the VSS-DLMF has superior steady-state performance as compared to the DLMF.

Throughout the paper, normal letters are used for scalars, boldface lowercase letters for vectors, and boldface uppercase letters for matrices. Moreover, $\mathbb{E}\{\cdot\}$ denotes expectation operation, $(\cdot)^\top$ represents transpose, $\|\cdot\|$ takes the ℓ_2-norm, $\mathrm{sgn}\{\cdot\}$ is the sign function, and ∇ is the gradient operator.

2. Signal model and DLMF

Fig. 1 illustrates a diffusion network consisting of N agents, where $u_k(i)$ and $d_k(i)$ are the input and observation signals at agent k, respectively, $\mathcal{N}_k$ represents the set of agents connected to agent k, including agent k itself. The observation signal of each agent k is normally given by the linear model

$$d_k(i) = \boldsymbol{u}_{k,i}^\top \boldsymbol{w}^o + v_k(i) \tag{1}$$

where $\boldsymbol{u}_{k,i} = [u_k(i), u_k(i-1), \ldots, u_k(i-M+1)]^\top$ is the input vector consisting of the M most recent samples of $u_k(i)$, $\boldsymbol{w}^o$ is an unknown vector to estimate, $v_k(i)$ denotes the measurement noise, which is assumed to be independent of any other signals in the network.

In [12], an adapt-then-combine (ATC) and a combine-then-adapt (CTA) diffusion strategies using a general convex cost function were proposed. We focus on the ATC diffusion strategy in this work, because it can achieve lower steady-state misalignment than the CTA diffusion strategy in some situations. Denote the estimate for $\boldsymbol{w}^o$ at agent k and time i by $\boldsymbol{w}_{k,i}$. The ATC diffusion strategy consists of the following adaptation and combination steps [12]:

$$\begin{cases} \boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} - \mu_k \sum_{l \in \mathcal{N}_k} c_{l,k} \nabla_{\boldsymbol{w}} J_l(\boldsymbol{w}_{k,i-1}) & \text{(2a)} \\ \boldsymbol{w}_{k,i} = \sum_{l \in \mathcal{N}_k} a_{l,k} \boldsymbol{\psi}_{l,i} & \text{(2b)} \end{cases}$$

where $\boldsymbol{\psi}_{k,i}$ is the intermediate estimate for $\boldsymbol{w}^o$ at agent k and time i, μ_k denotes a small positive step-size parameter, $J_l(\cdot)$ is a general convex cost function, $c_{l,k}$ and $a_{l,k}$ are the adaptation and combination weights of agent l on agent k, respectively, which satisfy

$$c_{l,k} \ge 0, \quad \sum_{k=1}^{N} c_{l,k} = 1, \text{ and } c_{l,k} = 0 \text{ if } l \notin \mathcal{N}_k \tag{3}$$

$$a_{l,k} \ge 0, \quad \sum_{l=1}^{N} a_{l,k} = 1, \text{ and } a_{l,k} = 0 \text{ if } l \notin \mathcal{N}_k. \tag{4}$$

There are several rules for selecting these weights, such as the uniform, maximum degree, Metropolis, relative degree, and relative degree-variance rules, which were summarized in [8].

Define the component cost function at agent l as

$$F_l(\boldsymbol{w}) = \mathbb{E}\left\{ \left[d_l(i) - \boldsymbol{u}_{l,i}^\top \boldsymbol{w} \right]^4 \right\}. \tag{5}$$

Obviously, this cost function is convex and therefore can be used in (2a). Its gradient with respect to $\boldsymbol{w}_{k,i-1}$ is

$$\nabla_{\boldsymbol{w}} F_l(\boldsymbol{w}_{k,i-1}) = -4\mathbb{E}\left\{ \boldsymbol{u}_{l,i} \left[d_l(i) - \boldsymbol{u}_{l,i}^\top \boldsymbol{w}_{k,i-1} \right]^3 \right\}. \tag{6}$$

Replacing $\nabla_{\boldsymbol{w}} J_l(\boldsymbol{w}_{k,i-1})$ in (2a) by the instantaneous value of $\nabla_{\boldsymbol{w}} F_l(\boldsymbol{w}_{k,i-1})$ in (6), we obtain

$$\begin{cases} \boldsymbol{\psi}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu_k \sum_{l \in \mathcal{N}_k} c_{l,k} \boldsymbol{u}_{l,i} \left[d_l(i) - \boldsymbol{u}_{l,i}^\top \boldsymbol{w}_{k,i-1} \right]^3 & \text{(7a)} \\ \boldsymbol{w}_{k,i} = \sum_{l \in \mathcal{N}_k} a_{l,k} \boldsymbol{\psi}_{l,i} & \text{(7b)} \end{cases}$$

which is the weight vector update equation of the DLMF. Note that the factor 4 appears in (6) is absorbed into the step-size μ_k in (7a).

3. Variable step-size development

The performance of the DLMF depends on the step-sizes $\{\mu_k\}$. If the step-sizes are large, then the DLMF converges fast, but its steady-state misalignment is high; if the step-sizes are small, then its steady-state misalignment is low, but it converges slowly. The DLMF needs to take a tradeoff between fast convergence rate and low steady-state misalignment by using constant step-sizes.

Research has verified that using a variable step-size can address the tradeoff problem existing in adaptive algorithms. The method of largest decrease of mean square deviation (MSD) is one of the popular methods to derive variable step-sizes [21]. However, this method does not apply to the LMF due to the cube of the error signals in its update equation. Since the DLMF is an extension of the LMF in network domain and therefore the method of largest decrease of MSD can also not be used to derive variable step-sizes for the DLMF, we need to use other methods to derive variable step-sizes.

In [22] a gradient descent method was proposed to derive a variable regularization parameter to address the tradeoff problem for the normalized LMS (NLMS). In [23] a normalized gradient descent method was presented to derive an improved variable regularization parameter. In the following, we will use the normalized gradient descent