Signal Processing 134 (2017) 268–274

Contents lists available at ScienceDirect

Signal Processing

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

Short communication

Diffusion least mean square/fourth algorithm for distributed estimation

Zongsheng Zheng, Zhigang Liu*, Ming Huang

School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China

ARTICLE INFO

Keywords:
Adaptive networks
Distributed estimation
Diffusion strategy
Mixed square/fourth error cost function

ABSTRACT

Proposed is a diffusion least mean square/fourth (LMS/F) algorithm, which is characterized by its fast convergence and low steady-state misalignment for distributed estimation in non-Gaussian noise environments. Instead of the conventional mean square error cost function, the diffusion LMS/F algorithm is derived from the mixed square/fourth error cost function, which is more suitable for non-Gaussian noise environments. Moreover, we incorporate the L_1- and L_0-norm constraints into the mixed square/fourth error cost function, and then a class of diffusion sparse LMS/F algorithms is developed which is able to exploit the sparsity of the considered system. Simulation results show that the diffusion LMS/F algorithm outperforms the conventional diffusion LMS and LMF algorithms in non-Gaussian noise environments. The improvements of diffusion sparse LMS/F algorithms in terms of steady-state misalignment are also demonstrated relative to the diffusion LMS/F algorithm.

1. Introduction

To estimate some parameters of interest from the data collected at nodes distributed over a geographic region, the distributed estimation was introduced [1–6]. In the distributed estimation, every node in the network communicates with a subset of the nodes, and the estimation is performed at each node in the network. There are a couple of distributed strategies that have been developed in the literature, namely, the incremental [1,6] and diffusion [2–5] strategies.

The diffusion strategy uses the subset of neighbors to communicate, and therefore requires low computational complexity and owns stable behavior in real-time adaptation. The diffusion least mean square (LMS) algorithm was first proposed in [2]. In [3], a general form of diffusion LMS algorithms was presented in which the adapt-then-combine (ATC) and combine-then-adapt (CTA) versions of diffusion LMS algorithms were formulated. The diffusion sparse LMS algorithms [4,5] were developed to enhance the detection of sparsity in the underlying system model. As presented in [3], the diffusion LMS algorithm is obtained based on the mean square error cost function. As is well-known, mean square error-based adaptive algorithms achieve optimal performance when the measurement noise is Gaussian. Besides, these mean square error-based adaptive algorithms show the same behavior for all noise distributions, since their performance depends only on the noise variance [7–9], which can also be seen from Fig. 3 in Section 5.1.

Research has shown that the adaptive algorithms based on high-order moment error cost function, e.g., least mean fourth (LMF) algorithm [7], performance better than the mean square error-based adaptive algorithms in some non-Gaussian noise environments, such as uniform or binary noise. However, the LMF algorithm has several stability problems that may put a limitation to its use in applications [8,9]. In addition, the LMF algorithm suffers from slow convergence in high signal-to-noise (SNR) environment [10–12]. Recently, by combining the benefit of the LMS and LMF algorithms, the least mean square/fourth (LMS/F) algorithm has been developed based on the mixed square/fourth error cost function [10–12]. The results showed that the LMS/F algorithm performs better than the conventional LMS and LMF algorithms. Motivated by the good performance of the LMS/F algorithm for non-Gaussian noise environments, we develop a diffusion LMS/F algorithm for distributed estimation in this work. In addition, the diffusion L_1-norm constraint LMS/F (L_1-LMS/F), diffusion reweighted L_1-norm constraint LMS/F (RL_1-LMS/F) and diffusion L_0-norm constraint LMS/F (L_0-LMS/F) algorithms are developed to exploit sparsity in the underlying system model. Simulations in non-Gaussian noise environments show that the diffusion LMS/F algorithm obtains better performance than the diffusion LMS and LMF algorithms, and the diffusion sparse LMS/F algorithms outperform the diffusion LMS/F algorithms for sparse system estimation.

2. Diffusion LMS algorithm

The considered network is composed of N nodes distributed over a geographic region. At each time instant i, each node k has access to the time realization $\{d_k(i), u_{k,i}\}$ of zero-mean random data $\{\boldsymbol{d}_k(i), \boldsymbol{u}_{k,i}\}$,

* Corresponding author.
E-mail addresses: bk20095185@my.swjtu.edu.cn (Z. Zheng), liuzg_cd@126.com (Z. Liu), bk20095186@my.swjtu.edu.cn (M. Huang).

http://dx.doi.org/10.1016/j.sigpro.2016.12.020
Received 28 August 2016; Received in revised form 16 December 2016; Accepted 19 December 2016
Available online 21 December 2016

where $d_k(i)$ is a scalar measurement and $u_{k,i}$ is a $1 \times M$ regression vector. The relationship between $d_k(i)$ and $u_{k,i}$ is supposed to be linear

$$d_k(i) = u_{k,i} w^o + v_k(i) \tag{1}$$

where w^o is the unknown M-dimensional parameter vector of interest and $v_k(i)$ is the measurement noise with variance $\sigma^2_{v,k}$. Here, it is assumed that $u_{k,i}$ and $v_k(i)$ are independent over time and space.

The diffusion LMS algorithm is obtained by minimizing the mean square error cost function for each node k[3]

$$J_k^{loc}(w) = \sum_{l \in N_k} c_{l,k} E[(\boldsymbol{d}_l(i) - \boldsymbol{u}_{l,i} w)^2] \tag{2}$$

where N_k denotes the set of nodes in the neighborhood of node k including itself, $\{c_{l,k}\}$ are real, non-negative, and satisfy $\sum_{l=1}^N c_{l,k} = 1$.

The diffusion LMS algorithm performs the estimation with two steps: adaptation and combination. According to the order of these two steps, the diffusion LMS algorithm is classified into the ATC and CTA diffusion LMS algorithms [3].

1) ATC diffusion LMS algorithm:

$$\begin{cases} \psi_{k,i} = w_{k,i-1} + \mu_k \sum_{l \in N_k} c_{l,k} u_{l,k}^T (d_l(i) - u_{l,i} w_{k,i-1}) \\ w_{k,i} = \sum_{l \in N_k} a_{l,k} \psi_{l,i} \end{cases} \tag{3}$$

2) CTA diffusion LMS algorithm

$$\begin{cases} \psi_{k,i-1} = \sum_{l \in N_k} a_{l,k} w_{l,i-1} \\ w_{k,i} = \psi_{k,i-1} + \mu_k \sum_{l \in N_k} c_{l,k} u_{l,k}^T (d_l(i) - u_{l,i} \psi_{k,i-1}) \end{cases} \tag{4}$$

where $\psi_{k,i}$ is an intermediate estimate of w^o at node k, μ_k is the step size, $\{a_{l,k}\}$ are real, non-negative, and satisfy $\sum_{l=1}^N a_{l,k} = 1$.

3. Diffusion LMS/F algorithm

For each node k, the mixed square/fourth error cost function is considered

$$J_k^{loc}(w) = \sum_{l \in N_k} c_{l,k} E[(\boldsymbol{d}_l(i) - \boldsymbol{u}_{l,i} w)^2 - \lambda \ln(\lambda + (\boldsymbol{d}_l(i) - \boldsymbol{u}_{l,i} w)^2)] \tag{5}$$

where λ is a positive design parameter.

Using the steepest-descent method, after the similar derivations as in [3], the diffusion LMS/F algorithm is obtained.

1) ATC diffusion LMS/F algorithm

$$\begin{cases} \psi_{k,i} = w_{k,i-1} + \mu_k \sum_{l \in N_k} c_{l,k} \frac{u_{l,i}^T (d_l(i) - u_{l,i} w_{k,i-1})^3}{\lambda + (d_l(i) - u_{l,i} w_{k,i-1})^2} \\ w_{k,i} = \sum_{l \in N_k} a_{l,k} \psi_{l,i} \end{cases} \tag{6}$$

2) CTA diffusion LMS/F algorithm

$$\begin{cases} \psi_{k,i-1} = \sum_{l \in N_k} a_{l,k} w_{l,i-1} \\ w_{k,i} = \psi_{k,i-1} + \mu_k \sum_{l \in N_k} c_{l,k} \frac{u_{l,i}^T (d_l(i) - u_{l,i} \psi_{k,i-1})^3}{\lambda + (d_l(i) - u_{l,i} \psi_{k,i-1})^2} \end{cases} \tag{7}$$

Remark #1. : The comparison of the mixed square/fourth cost function with the square cost function and fourth-order cost function is shown in Fig. 1. From this figure, we observe that the mixed square/fourth cost function exhibits comparable steepness as the fourth-order cost function for small perturbations of the error, and the mixed square/fourth cost function exhibits comparable steepness as the square cost function for relatively larger error values. Hence, by using the mixed square/fourth cost function, the proposed algorithm obtains comparable convergence rate with the conventional diffusion LMS algorithm and achieves lower steady-state misalignment through the use of the fourth-order statistics for small perturbations of the error.

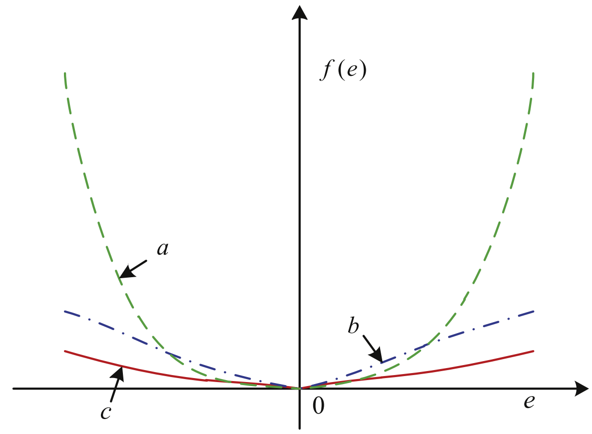

Fig. 1. Comparison of different stochastic cost functions, (a) fourth-order cost function $f(e) = e^4$, (b) square cost function $f(e) = e^2$, (c) mixed square/fourth cost function $f(e) = e^2 - \lambda \ln(\lambda + e^2)$, $\lambda = 1$.

Remark #2. : In [11], the stability of the LMS/F algorithm has been discussed. It reveals that the LMS/F algorithm owns comparable stability with the LMS algorithm. Moreover, as illustrated in Fig. 1, the mixed square/fourth cost function exhibits comparable steepness as the square cost function for relatively larger error values and, consequently, the proposed algorithm inherits the stability of the diffusion LMS algorithm.

4. Extend to sparse distributed estimation

The diffusion LMS/F algorithm is extended to sparse distributed estimation in this section. To seek an estimate of w^o at node k, we minimize the following penalized mixed square/fourth error cost function with L_p-norm

$$J_k^{loc}(w) = \sum_{l \in N_k} c_{l,k} E[(\boldsymbol{d}_l(i) - \boldsymbol{u}_{l,i} w)^2 - \lambda \ln(\lambda + (\boldsymbol{d}_l(i) - \boldsymbol{u}_{l,i} w)^2)] + \gamma \|w\|_p \tag{8}$$

where $\|w\|_p$ denotes the L_p-norm with respect to w, and γ is the weight given to the L_p norm constraint.

Following the derivation in [5], we obtain the diffusion sparse LMS/F algorithm below.

1) ATC diffusion sparse LMS/F algorithm

$$\begin{cases} \psi_{k,i} = w_{k,i-1} + \mu_k \sum_{l \in N_k} c_{l,k} \frac{u_{l,i}^T (d_l(i) - u_{l,i} w_{k,i-1})^3}{\lambda + (d_l(i) - u_{l,i} w_{k,i-1})^2} - \mu_k \gamma \xi_p(w_{k,i-1}) \\ w_{k,i} = \sum_{l \in N_k} a_{l,k} \psi_{l,i} \end{cases} \tag{9}$$

2) CTA diffusion sparse LMS/F algorithm

$$\begin{cases} \psi_{k,i-1} = \sum_{l \in N_k} a_{l,k} w_{l,i-1} \\ w_{k,i} = \psi_{k,i-1} + \mu_k \sum_{l \in N_k} c_{l,k} \frac{u_{l,i}^T (d_l(i) - u_{l,i} \psi_{k,i-1})^3}{\lambda + (d_l(i) - u_{l,i} \psi_{k,i-1})^2} - \mu_k \gamma \xi_p(\psi_{k,i-1}) \end{cases} \tag{10}$$

where $\xi_p(w) = [\xi_p(w_1), \xi_p(w_2), \dots \xi_p(w_m) \dots , \xi_p(w_M)]$ is the derivative of the L_p-norm with respect to w, and w_m denotes the mth component of w. Next, we will describe $\xi_p(w_m)$ for each diffusion sparse LMS/F algorithm in detail.

1) Diffusion L_1-LMS/F algorithm

$$\xi_1(w_m) = \text{sgn}(w_m) = \begin{cases} \frac{w_m}{|w_m|}, & w_m \neq 0 \\ 0, & w_m = 0 \end{cases} \tag{11}$$

2) Diffusion RL_1-LMS/F algorithm