Signal Processing 149 (2018) 124–134

Contents lists available at ScienceDirect

Signal Processing

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

Complex-Valued adaptive networks based on entropy estimation

Gang Wang[a], Rui Xue[b,*], Chao Zhou[c], Junjie Gong[c]

[a] Center for Robotics, School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, PR China

[b] School of Electronics and Information Engineering, Beihang University, Beijing, PR China

[c] School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, PR China

ARTICLE INFO

Article history:
Received 21 July 2017
Revised 26 February 2018
Accepted 10 March 2018
Available online 14 March 2018

Keywords:
Gaussian entropy
Multi-sensor network (MSN)
Mean-square error (MSE)
Degree of non-circularity (DNC)

ABSTRACT

In distributed estimation, the mean-square error (MSE) criterion has been extensively studied. When complex-valued signals are involved, the additive noise can present different degrees of non-circular properties. The MSE criterion can be optimal only when the error signal is circular, and may not perform well for non-circular error signal. To improve the performance, we present a new diffusion adaptive strategy using the Gaussian entropy criterion as the cost function. Complex-valued Gaussian entropy was early introduced for linear and widely linear filtering. Unfortunately, the closed-form solution based on Gaussian entropy was not obtained due to the nonlinearity of the entropy equation. In this paper, we derive a closed-form solution based on Gaussian entropy for linear and widely linear filters, and provide mean value steady and mean-square performance analysis for the network in detail. Our theoretical analysis demonstrates that the steady-state error approaches zero when the additive noise is maximally non-circular. The simulations demonstrate that the proposed method outperforms the MSE criterion for non-circular noise.

1. Introduction

Distributed estimation derives from good examples of real-time learning and adaptation behavior, and it plays a key role in in-network signal processing. Compared with centralized estimation, it only uses local data cooperation between neighbor nodes to increase robustness and reduce complexity [1–13]. Distributed estimation is commonly used in many contexts, such as the foraging behavior of animals [14], distributed detection [15] and target tracking and escaping from predators [16]. The algorithms based on different cooperation rules include diffusion least mean square (LMS) [1–17], asynchronous diffusion adaptation [18], incremental LMS [19], diffusion Kalman filtering [20,21], diffusion recursive least square (RLS) [22], distributed sparse RLS [23], diffusion augmented complex LMS for non-circular complex signals [24], minimization of complex-error entropy for complex-valued filtering [25] and diffusion information theoretic learning [26]. Among the algorithms, the mean-square error (MSE) criterion is commonly used.

When complex-valued signals are involved in real-world applications, they can present different degrees of non-circularity (DNC) properties. Complex-valued signals are typically characterized by their second-order statistical (SOS) properties [27–35]. The SOS is represented by the conventional covariance matrix and the pseudo-covariance matrix, which characterizes the improperness of the signal. The MSE criterion is optimal only when the error signal is circular, and may perform poorly for non-circular error signal. To overcome this drawback, complex-valued Gaussian entropy was introduced in [31,32] for linear and widely linear filtering, and the DNC of the error can be taken into account.

It was shown in [25] that the performance of the network depends on the DNC of the complex signal. This motivates us to introduce the Gaussian entropy criterion into the complex-valued network.

The simulations in [31] demonstrated the following:

1) for a linear filter, the entropy criteria always outperform the MSE with non-circular additive noise; and
2) for a widely linear filter, the MSE and entropy criteria have the same performance with non-circular additive noise.

However, there was no theoretical explanation for the two phenomena because a closed-form solution based on Gaussian entropy was not obtained in [31] due to the nonlinearity of the entropy equation. If a closed-form solution is given, we can apply the entropy criteria to the network robustly.

Since it is difficult to obtain the closed-form solution of Gaussian entropy, we turn to solve the real and imaginary parts of the closed-form solution. In this paper, we derive a closed-form solu-

* Corresponding author.
E-mail addresses: wanggang_hld@uestc.edu.cn (G. Wang), xuerui@buaa.edu.cn (R. Xue), gong_junjie34@163.com (J. Gong).

https://doi.org/10.1016/j.sigpro.2018.03.003
0165-1684/

tion based on Gaussian entropy for the linear and widely linear filters, and determine the following:

1) for a linear filter, the mean-square deviance (MSD) decreases with an increase of the DNC of additive noise under the same signal-to-noise ratio (SNR), and approaches zero when all additive noise is maximally non-circular under different SNRs; and
2) for a widely linear filter, the solution based on Gaussian entropy is equivalent to the solution based on the MSE.

Then we introduce complex-valued Gaussian entropy as the cost function for the multi-sensor network to process non-circular noise. The Gaussian entropy performance's dependence on the DNC is studied. The DNC is the absolute value of the circularity coefficient (CC) [31], which works in the gradient of Gaussian entropy. The selection of the CC plays a key role in the performance of the iteration. The estimation of the CC is introduced during the iterations. It can be expected that Gaussian entropy criteria can achieve better performance than the traditional LMS algorithm for a network with non-circular noise.

The paper is organized as follows: In Section 2, the problem statement is explained, including two types of gradient process. In Section 3, the closed-form solution for Gaussian entropy is proposed. In Section 4, a novel distributed adaptive methodology is presented that relies on cooperative diffusion protocols. In Section 5, the mean value stability and mean-square performance of the integrated adaptive network are analyzed. The simulation results are presented in Section 6 and the conclusion is presented in Section 7.

To help the analysis, the superscript * denotes a complex conjugate, and superscripts T and H denote the transpose and conjugate transpose, respectively.

2. Problem statement

For a zero mean complex random variable $x = x_R + jx_I$, we have the following definitions.

Definition 1. (*Simply circular*) If x and $x\exp(j\theta)$ have the same probability density function (PDF) for $-\pi \le \theta < \pi$, then the complex random variable x is called simply circular or strictly circular.

Definition 2. (*Second-order circular*) If $E\{x^2\} = 0$, where x is zero mean, then x is called second-order circular.

Definition 3. (*Degree of non-circularity (DNC)*) The DNC of a complex random variable x is defined as the index $|\rho|$ [31]:

$$|\rho| = \frac{\left|E\left\{x^2\right\}\right|}{E\left\{|x|^2\right\}}, \tag{1}$$

where x is zero mean and $0 \le |\rho| \le 1$.

Definition 4. (*Circularity coefficient (CC)*) The CC of the complex random variable x is defined as

$$\rho = \frac{E\left\{x^2\right\}}{E\left\{|x|^2\right\}}, \tag{2}$$

where x is zero mean and $0 \le |\rho| \le 1$.

Remark 1. The CC ρ equals zero for a circular signal and differs from zero for a non-circular signal. For Gaussian processes, the second-order circularity implies strict circularity, and a zero mean random vector x is second-order circular when $E\{xx^T\} = 0$.

Consider a network of N nodes distributed over a spatial region (see Fig. 1). Each node k has access to its time series $\{d_k(i), \boldsymbol{u}_k(i)\}$, $k = 1, 2, \ldots, N$. Two nodes are neighbors when they can share data. Let N_k denote the neighborhood of node k. There is an unknown vector $\boldsymbol{w}^o \in \mathbb{C}^M$ at each node k to estimate the temporal wide-sense stationary measurements. $d_k(i)$ and $\boldsymbol{u}_k(i)$ over time $i \ge 0$ are assumed to obey the traditional model [6]:

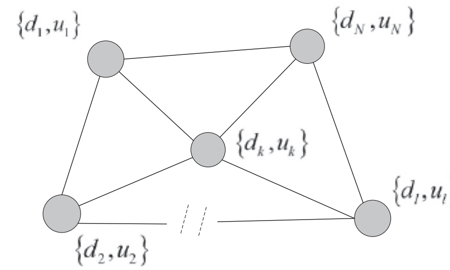

Fig. 1. Distributed network with N active nodes accessing space-time data.

$$\begin{aligned} d_k(i) &= \boldsymbol{u}_k(i)\boldsymbol{w}^o + v_k(i), \\ \boldsymbol{u}_k(i) &= [u_k(i)\ u_k(i-1) \cdots u_k(i-M+1)], \\ \boldsymbol{w}^o &= [w_1 w_2 \cdots w_M]^T, \end{aligned} \tag{3}$$

where $d_k(i) \in \mathbb{C}^1$ is the desired signal, $\boldsymbol{u}_k(i) \in \mathbb{C}^{1\times M}$ is the regression row vector and $v_k(i)$ is zero mean background noise with variance $\sigma_{v,k}^2$.

In this research, the following assumptions are made:

A1) Additive noise is a white process, that is, $E\{v_k(i)v_l(j)\} = E\{v_k(i)v_l^*(j)\} = \begin{cases} 0, & k \ne l \\ 0, & k = l, j \ne i \end{cases}$
A2) The inputs of different nodes $\boldsymbol{u}_k(i)$ are spatially and temporally independent, that is $E\{\boldsymbol{u}_k(i)\boldsymbol{u}_l^H(j)\} = 0$ for $j \ne i$ and $k \ne l$.
A3) The inputs of all the nodes and additive noise are uncorrelated, that is, $E\{\boldsymbol{u}_k(i)v_l(j)\} = E\{\boldsymbol{u}_k^*(i)v_l(j)\} = 0$.

The above assumptions are customary in the context of adaptive filters [3,31], and are in fact required to simplify the analysis. It is important to highlight that, although assumptions A1 and A3 can generally be considered strong in practice, assumption A2 is often weak particularly in networks where the sensors are close to each other. This characteristic must be carefully taken into account when the proposed expressions are used for analysis purposes.

2.1. MSE criterion

When noise $v_k(i)$ is circular, the unknown parameter of interest $\boldsymbol{w}^o$ can be estimated by minimizing the MSE cost function [1,5]:

$$\begin{aligned} J_{MSE}(\boldsymbol{w}) &= \sum_{k=1}^{N} E|e_k(i)|^2 = \sum_{k=1}^{N} E\left\{e_k(i)e_k^*(i)\right\}, \\ e_k(i) &= d_k(i) - \boldsymbol{u}_k(i)\boldsymbol{w}_k(i), \end{aligned} \tag{4}$$

where $\boldsymbol{w}_k(i)$ is the estimate for $\boldsymbol{w}^o$ at node k at time i. The gradient of the MSE is given by

$$\frac{\partial J_{LMS}}{\partial \boldsymbol{w}^*} = -\sum_{k=1}^{N} E\left\{e_k \boldsymbol{u}_k^H\right\}. \tag{5}$$

At the stationary point, we have $\sum_{k=1}^{N} E\{e_k\boldsymbol{u}_k^H\}=0$ and the optimal solution [3]:

$$\begin{aligned} \boldsymbol{w}^o &= \left[\sum_{k=1}^{N} \boldsymbol{R}_{u,k}\right]^{-1}\left[\sum_{k=1}^{N} \boldsymbol{R}_{du,k}\right], \\ \boldsymbol{R}_{u,k} &= E\left\{\boldsymbol{u}_k^H(i)\boldsymbol{u}_k(i)\right\}, \\ \boldsymbol{R}_{du,k} &= E\left\{d_{k,i}\boldsymbol{u}_k^H(i)\right\}. \end{aligned} \tag{6}$$

A simple adaptive learning algorithm for the estimation for $\boldsymbol{w}^o$ at node k at time i is given by the LMS algorithm [8]:

$$\boldsymbol{w}_k(i) = \boldsymbol{w}_k(i-1) + \mu_k e_k(i)\boldsymbol{u}_k^H(i-1) \tag{7}$$

where μ_k is the step size at node k.