Signal Processing 150 (2018) 166–170

Contents lists available at ScienceDirect

Signal Processing

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

Short communication

Switching criterion for sub-and super-Gaussian additive noise in adaptive filtering

Gang Wang[a], Rui Xue[b,*], Ji Zhao[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 Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, PR China

ARTICLE INFO

Article history:
Received 27 February 2018
Revised 10 April 2018
Accepted 18 April 2018
Available online 19 April 2018

Index Terms:
Adaptive filter
Least mean square
Maximum correntropy criterion
Least mean fourth
Normalized kurtosis

ABSTRACT

Additive noise distributions can be divided into three types: Gaussian, super- and sub-Gaussian. The existing algorithms for adaptive filtering do not provide a better performance than the least mean square (LMS) method for the super- and sub-Gaussian noise simultaneously. For example, the maximum correntropy criterion performs better (worse) than the LMS method for super-Gaussian (sub-Gaussian) noise, whereas the least mean fourth performs better (worse) than the LMS method for sub-Gaussian (super-Gaussian) noise. We propose a criterion for switching between sub- and super- Gaussian additive noise, that could be used to assess whether the error signal had a sub- or super-Gaussian profile, and thus determine which algorithm would work best in the iterative process. Simulations demonstrate that the switching criterion helps the proposed switching algorithm to produce a better performance than the LMS algorithm for sub and super-Gaussian noise simultaneously.

1. Introduction

Adaptive filtering is widely used in signal processing, and its most popular cost function is the mean square error (MSE) [1–3]. The corresponding stochastic gradient descent algorithm is generally called least mean square (LMS). The LMS algorithm relies on second-order similarity measures, and performs well with Gaussian additive noise, for which the higher statistics are constants. Recently, information theoretic learning (ITL) [4–20] has been proposed for uses with the higher-order moments as a cost function and may work well for non-Gaussian noise, for which the higher statistics are not constants. The available cost functions include the maximum correntropy criterion (MCC) [4–12], the improved least sum of exponentials (ILSE) [13], the least mean kurtosis (LMK) [14] and the least mean fourth (LMF) [15–20].

Generally speaking, the distributions of the signal are divided into three types [11,21,22]: Gaussian, super-Gaussian, and sub-Gaussian distribution. Sub- and super-Gaussian signals are both non-Gaussian signals. Typical sub-Gaussian noise has a uniform distribution, while typical super-Gaussian noise is impulse noise.

In an impulsive noise environment, the MCC algorithm outperforms the LMS [4,20]. In a Gaussian noise environment, the LMS algorithm performs better than the LMF algorithm [15] and the MCC algorithm yields a smaller steady state excess mean square error (EMSE) than the LMS algorithm with the same step-size, while as the kernel width increases, the steady-state EMSE of the MCC will approach that of the LMS [8]. In a uniform additive noise environment, the LMF algorithm outperforms the LMS algorithm [15,20]. Typically, the LMS, LMF and MCC algorithms work well for Gaussian, sub-Gaussian, and super-Gaussian noise, respectively.

To our knowledge, the LMF and LMS have not been compared in the impulsive noise environment, and the MCC and LMS have not been compared in the uniform noise environment. Our simulations show that most of the existing algorithms (the MCC, ILSE, LMF, and LMK) cannot perform better than the LMS for sub- and super-Gaussian noise simultaneously. For super-Gaussian noise, a steady error comparison under the same initial convergence speed *roughly* ranks the algorithms in the order MCC < LMS < LMF. For sub-Gaussian noise, the same error comparison *roughly* ranks the algorithms in the order LMF < LMS < MCC. It should be mentioned here that the rank is not rigorous, and the rank is currently being explored.

Early independent component analysis (ICA) algorithms [21,22] introduced the normalized kurtosis to access the sub- and super-Gaussian signals. This motivate us to proposed a

* Corresponding author.
E-mail addresses: wanggang_hld@uestc.edu.cn (G. Wang), xuerui@buaa.edu.cn (R. Xue), zhaoji@std.uestc.edu.cn (J. Zhao).

https://doi.org/10.1016/j.sigpro.2018.04.017
0165-1684/

switching criterion to assess whether the error signal has sub- or super-Gaussian distribution, and can thus determine which algorithm (MCC or LMF) will work best in the iterative processes. Moreover, the following cost function, denoted by $G(x)$, was often used to separate non-Gaussian signals in ICA algorithms [21–28],

$$G(x)=\frac{1}{\alpha}\log\cosh(\alpha x), \tag{1}$$

where $1 \le \alpha \le 2$ and $\alpha = 1$ is often adopted for that this parameter affects the performance of the proposed algorithm slightly. Compared with the MCC algorithm, which depends on the different kernel parameter for different impulse noise, the parameter αin $G(x)$ can be fixed. Since the cost function of MCC and $G(x)$ are both used in ICA algorithm to separate non-Gaussian signals, we introduce $G(x)$ to adaptive filtering applications, and believe that it would work well for super-Gaussian noise, and use this cost function to substitute for the MCC algorithm in the switching criterion.

The contributions of this paper can be summarized as follows:

1) We propose the use of normalized kurtosis as a switching criterion to assess whether the additive noise is sub-Gaussian or super-Gaussian in adaptive filtering.
2) We propose a new cost function, $G(x)$, for processing of super-Gaussian noise.

The paper is organized as follows: In Section 2, the problem is stated in detail. In Section 3, the new cost function is introduced for super-Gaussian noise. In Section 4, a switching criterion for processing sub or super-Gaussian noise is presented. In Section 5, simulations are described and results are provided. Finally, a conclusion is drawn in Section 6.

2. Problem statement

The normalized kurtosis [21,22] can be considered as a measure of the non-Gaussianity of the error signal. The normalized kurtosis of a random variable x is defined as

$$\kappa_4 = \frac{E\left\{x^4\right\}}{E^2\{x^2\}} - 3 \tag{2}$$

where x is zero mean with unit variance.

A distribution with negative normalized kurtosis is then called sub-Gaussian, or short-tailed (e.g., uniform). A distribution with positive normalized kurtosis is called super-Gaussian, or heavy-tailed (e.g., Laplacian). A zero-kurtosis distribution is called Gaussian [21,22].

When a linear filtering problem is considered, there is an input vector $\boldsymbol{u} \in \mathbb{R}^M$, with unknown parameter $\boldsymbol{w}_o \in \mathbb{R}^M$ and the desired response $d \in \mathbb{R}^1$. Data $d(i)$ are observed at each time point i using the linear regression model:

$$d(i) = \boldsymbol{w}_o^T \boldsymbol{u}(i) + v(i), \quad i = 1, 2, \ldots, L \tag{3}$$

where v is the zero mean background noise with variance σ_v^2 and L is the sequence length. The error signal for the linear filter is defined as

$$e(i) = d(i) - \boldsymbol{w}^T \boldsymbol{u}(i) \tag{4}$$

where $\boldsymbol{w}$ is the estimated value of $\boldsymbol{w}_o$.

The linear filtering algorithms of the LMS, MCC, and LMF are as follows. The cost function based on the MSE is given by

$$J_{MSE}(\boldsymbol{w}) = E\left\{e^2\right\} \tag{5}$$

where E denotes the expectation operator. The corresponding stochastic gradient descent or LMS algorithm is

$$\boldsymbol{w}_{LMS}(i+1) = \boldsymbol{w}_{LMS}(i) + \mu e(i)\boldsymbol{u}(i) \tag{6}$$

where μ denotes the step size and $\mu > 0$.

The cost function based on the correntropy of the error, which is also known as the MCC, is given by

$$J_{MCC}(\boldsymbol{w}) = E\left\{\exp\left(-\frac{e^2}{2\sigma^2}\right)\right\} \tag{7}$$

where σdenotes the kernel bandwidth. The corresponding stochastic gradient ascent algorithm is

$$\boldsymbol{w}_{MCC}(i+1) = \boldsymbol{w}_{MCC}(i) + \mu \exp\left(-\frac{e^2(i)}{2\sigma^2}\right) e(i)\boldsymbol{u}(i). \tag{8}$$

The cost function based on the LMF is given by

$$J_{LMF}(\boldsymbol{w}) = E\left\{e^4\right\}. \tag{9}$$

The corresponding stochastic gradient descent algorithm is

$$\boldsymbol{w}_{LMF}(i+1) = \boldsymbol{w}_{LMF}(i) + \mu e^3(i)\boldsymbol{u}(i). \tag{10}$$

3. Proposed cost function for super-Gaussian noise

In practice, the convergence speed and steady error of the MCC algorithm depend on the kernel bandwidth. A fixed kernel bandwidth may not work well for different super-Gaussian noise. The correct choice of kernel width in the MCC algorithm imposes a trade-off among robustness, convergence rate and steady-state accuracy. The adaptive kernel width MCC algorithms [5–7] can improve the learning speed especially when the initial weight vector is far away from the optimal weight vector.

We propose another cost function for linear filtering as follows:

$$G(\boldsymbol{w})=\frac{1}{\alpha}\log\cosh[\alpha e(i)] \tag{11}$$

where $1 \le \alpha \le 2$, and $\alpha = 1$ in the simulations. The derivative of (11) is given by

$$g(\boldsymbol{w})=-\tanh[\alpha e(i)]\boldsymbol{u}. \tag{12}$$

The corresponding stochastic gradient descent algorithm is

$$\boldsymbol{w}_G(i+1) = \boldsymbol{w}_G(i) + \mu\tanh[\alpha e(i)]\boldsymbol{u}(i). \tag{13}$$

In the ICA algorithm [28], Eq. (7) and (11) are both used in the fast-ICA algorithm to separate the super-Gaussian sources, and Eq. (9) was used to separate sub-Gaussian sources. Eq. (7) is robust to highly super-Gaussian signals.

4. Switching criterion for the sub and super-Gaussian noise

We are unlikely to know in advance whether the additive noise is sub- or super-Gaussian. The MCC and the LMF do not perform better than the LMS algorithm for sub-Gaussian and super-Gaussian noise simultaneously. Using the normalized kurtosis, we propose a switching criterion for the sub- and super-Gaussian error signals.

The normalized kurtosis κ_4 can be estimated using the following adaptive form [21]:

$$\begin{aligned} M_4(i+1) &= (1-\eta)M_4(i) + \eta e^4(i), \\ M_2(i+1) &= (1-\eta)M_2(i) + \eta e^2(i), \\ \hat{\kappa}_4(i+1) &= M_4/(M_2)^2 - 3, \end{aligned} \tag{14}$$

where $\hat{\kappa}_4$ represents the estimate of κ_4, η denotes the forgetting factor, and $0 < \eta < 1$.

The switching algorithm, which combines the MCC, LMS with LMF algorithm, is given by