Robust Diffusion Recursive Adaptive Filtering Algorithm Based on $l_p$-norm

MA Wentao¹, DUAN Jiandong¹, GUI Guan², CHEN Badong³

1. Department of Electrical Engineering, Xi’an University of Technology, Xi’an 710048, China
   E-mail: xjtu.wentao@gmail.com, Duanjd@xaut.edu.cn
2. College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
   E-mail: guiguan@njupt.edu.cn
3. School of Electronic and Information Engineering, Xi’an Jiaotong University, Xi’an, 710049, China
   E-mail: chenbd@mail.xjtu.edu.cn

Abstract: A robust diffusion adaptive filtering algorithm, called the diffusion recursive least $l_p$-norm (DRLP), is developed for distributed estimation over network. The new algorithm aims at recursively minimizing the $l_p$-norm of error, and can offer a more stable and robust solution than traditional adaptive filtering schemes based on minimization of the squared error, such as the diffusion recursive least squares (DRLS) algorithm. Simulation results show that the proposed DRLP can outperform several state-of-the-art methods especially when the network is disturbed by impulsive noises.

Key Words: Diffusion, $l_p$-norm, Distributed network, Robust, Impulsive noise

1 Introduction

Most existing diffusion adaptive filtering algorithms for distributed estimation over network, such as the diffusion least mean square (DLMS) [1-2], the diffusion recursive least squares (DRLS)[3], and many variants [4-6], are derived by minimizing the $l_2$-norm of error. These algorithms usually perform very well in Gaussian noises. However, they suffer convergence and stability problems especially when the underlying system is contaminated with heavy-tailed non-Gaussian noises (or impulsive noises), such as the $\alpha$ -stable noise [7], whose second order moments with $\alpha < 2$ are infinite. To address this issue, some new diffusion adaptive filtering algorithms based on the $l_p$-norm [8] or information theoretic learning (ITL) criteria [9] are developed. $l_p$-norm can be as a robust criterion for adaptive filter, and it has been used in [10-12] to develop different types of adaptive algorithm. The diffusion least mean $l_p$-power (DLMP) based on $l_p$-norm error criterion was proposed to estimate the parameters of the wireless sensor networks [13]. For non-Gaussian cases, information theoretic learning provides a more general framework and can also achieve a desirable performance. The diffusion adaptive filter algorithm employing minimum error entropy[14] was proposed in [15]. Under the minimum error entropy criterion, the entropy of a batch of N recent most error samples is used as a cost function to be minimized to adapt the weights. The evaluation of the error entropy involves a double sum over the samples, which is computationally expensive especially when

*This work is supported by 973 Program (No. 2015CB351703) and the National Natural Science Foundation of China (no. 61372152).
estimate an unknown deterministic column vector \( w_o \) of size \( M \times 1 \). At each instant \( i \) (\( i = 1, 2, \cdots \)), each node \( k \) has access to the realizations of a scalar measurement \( d_k^i(i) \) and a regression vector \( u_{kj} \) of size \( 1 \times M \), related via
\[
d_k^i(i) = u_{kj}^T w_o + n_k(i)
\]
where \( n_k(i) \) denotes the measurement noise with impulsive character mainly considered in this work. Our objective is to develop a diffusion RLP algorithm, which shows the robustness of dealing with impulsive noise and consequently, gives a good estimate for \( w_o \).

2.2 Diffusion Recursive Least \( l_p \)-norm algorithm

The objective of the proposed DRLP algorithm is to estimate \( w_o \) by solving the following weighted least \( l_p \)-norm problem
\[
w_j^i = \arg \min_{w} \sum_{j=1}^{i} \lambda^{i-j} \sum_{l=1}^{N} |d_j^i(j) - u_{lj}^i w|^p
\]
where \( \lambda \) denotes a forgetting factor with \( 0 < \lambda \leq 1 \).

We refer to this problem as the global least \( l_p \)-norm problem, since at time \( i \), the solution takes into account all measurements from all nodes up to time \( i \). We now proceed to propose a diffusion estimation scheme where each node has only access to limited data, namely, the data from the neighbouring nodes. Thus, the node \( k \) can locally solve the following least \( l_p \)-norm problem
\[
\psi_{k,j} = \arg \min_{w} \sum_{j=1}^{i} \lambda^{i-j} \sum_{l=1}^{N} c_{l,k} |d_j^i(j) - u_{lj}^i w|^p
\]
where \( c_{l,k} \) is a positive weighting coefficient representing the cooperation rule among nodes, satisfying
\[
c_{l,k} = 0 \text{ if } l \notin N_k \text{ and } \sum_{l=1}^{N} c_{l,k} = 1.
\]

The vector \( \psi_{k,j} \) denotes a local estimate at node \( k \). One can obtain an estimate at node \( k \) by a weighted average of the local estimates in the neighbourhood of node \( k \), which takes the form
\[
w_{k,j} = \sum_{l \in N_k} a_{l,k} \psi_{l,j}
\]
with \( a_{l,k} \) just being like the coefficient \( c_{l,k} \). Note that the nodes in the network will generally have access to different data. So their estimates of \( w_o \) will be solutions to different least \( p \)-power problems. Naturally, we want these estimates to be close to the global least \( l_p \)-norm solution of (2).

Remark 1: The equation (3) will be the least squares problem with non-regularized proposed in [3] when \( p = 2 \).

According to \( l_p \)-norm aware (including RLP and LMP) algorithms, we know that the term only \( |e_i^j(i)|^{p-2} \) is added comparing with the \( l_2 \)-norm aware (including RLS and LMS) algorithms when \( p \neq 2 \). This term can be as a step size for \( l_p \)-norm aware algorithms. When these algorithms have not converged, the \( |e_i^j(i)| \) is large. This makes \( p \)-norm aware algorithms have fast convergence speed. When these algorithms have converged, the error \( |e_i^j(i)| \) is small. This makes the misadjustment be small. This character makes the \( l_p \)-norm aware algorithms very suitable for addressing the signal processing case under impulsive noise.

For dealing with distributed estimation over network under impulsive noise environment, we propose the DRLP algorithm based on (3) to collectively estimate \( w_o \) from individual measurements in two steps (including an incremental update and a spatial update stage) as follows:

i) At time \( i \), the nodes communicate their measurements \( d_k^i(i) \) and regressors \( u_{kj} \) with their neighbors, and use these data to update their local estimates using RLP iterations (named incremental update). The resulting preestimates are named \( \psi_{k,j} \) as in (3).

ii) The nodes communicate their local preestimates with their neighbors and perform a weighted average to obtain the estimate \( w_{k,j} \) (named spatial update).

The above update stage is similar to the DRLS [3]. The DRLP algorithm is shown schematically in Table 1, and the detailed derivation of this algorithm is omitted in this paper which is similar to the DRLS algorithm.

**Table 1**: Diffusion Recursive Least \( l_p \)-norm Algorithm

**ALGORITHM**

Step 1: Initialization: \( w_{k,-1} = 0 \) and \( P_{k,-1} = \theta^{-1} I_M \), where \( P_{k,-1} \) is an \( M \times M \) matrix and \( \theta > 0 \), and \( I_M \) denotes an identity matrix.

Step 2: Incremental update

For every node \( k \), repeat
\[
\psi_{k,j} \leftarrow w_{k,j-1}, \quad P_{k,j} \leftarrow \lambda^{-1} P_{k,j-1}
\]

For all \( l \in N_k \)
\[
K_{k,l} = \frac{P_{k,l} u_{lj}}{\lambda + f(e_l(i)) c_{l,k} u_{lj}^T P_{k,l} u_{lj}}
\]
\[
\psi_{k,l} \leftarrow \psi_{k,l} + f(e_l(i)) c_{l,k} K_{k,l} (d_l(i) - u_{lj} \psi_{k,l})
\]
\[
P_{k,l} \leftarrow P_{k,l} - P_{k,l} f(e_l(i)) c_{l,k} K_{k,l} u_{lj} P_{k,l}
\]
\[
f(e_l(i)) = |e_l(i)|^{p-2} |d_l(i) - u_{lj} \psi_{k,l}|^{p-2}
\]

end

Step 3: Spatial update
\[
w_{k,j} = \sum_{l \in N_k} a_{l,k} \psi_{l,j}
\]

Remark 2: The DRLP differs from the conventional DRLS in that the nonlinearity of error will be included in the incremental update stage, and it means that the complexity of
the DRLP is more than one extra nonlinearity computation compared with DRLS [3]. In fact, the DRLP is a generalized version of the DRLS, which will reduce to the standard DRLS when $p = 2$.

3 Simulation Results

We conduct simulations to demonstrate the performance of the proposed DRLP algorithm. The topology of the network with 20 nodes is generated as a realization of the random geometric graph model as shown in Fig. 1. The location coordinates of the agents in the square region $[0,1.2] \times [0,1.2]$. The unknown parameter vector is set to $(M,1)^\text{stable}$, where $\text{randn}(\cdot)$ is the function of generating Gaussian random. The input regressors are zero-mean Gaussian, independent in time and space with size $M = 5$. In this paper, we consider the measurement noise as an alpha-stable noise. The $\alpha$-stable distribution provides a good model for such heavy-tailed noises [7]. The characteristic function of the alpha-stable process is given by

$$f(t) = \exp\left\{ j\delta t - \gamma |t|^\alpha \left[ 1 + j\beta \text{sgn}(t) S(t,\alpha) \right] \right\}$$

in which

$$S(t,\alpha) = \begin{cases} \frac{\alpha \pi}{2} & \text{if } \alpha \neq 1 \\ \frac{2}{\pi} \log |t| & \text{if } \alpha = 1 \end{cases}$$

where $\alpha \in (0, 2]$ is the characteristic factor, $-\infty < \delta < +\infty$ is the location parameter, $\beta \in [-1,1]$ is the symmetry parameter, and $\gamma > 0$ is the dispersion parameter. The characteristic factor $\alpha$ measures the tail heaviness of the distribution. The parameters vector of the noise model is defined as $V_{\alpha\text{-stable}}(\alpha, \beta, \gamma, \delta)$.

In our simulation, we mainly compare the performance of the proposed DRLP with the DRLS, DLMP and DMCC. All parameters of these algorithms are set by scanning for the best results. We set the parameter vector $V_{\alpha\text{-stable}}(1.2, 0, 1, 0)$ to generate the impulsive noise.

Fig. 2 gives the convergence curves in terms of MSD. We set the $p = 1$ for DLMP and DRLP algorithms. One can observe that DLMP, DMCC and DRLP work well when large outliers occur, while DRLS fluctuate dramatically due to the sensitivity to the impulsive noises. As can be seen from this result, the proposed DRLP algorithm has excellent performance in convergence rate and accuracy compared with other methods, and exhibits a significant improvement in robust performance under the impulsive noise environment.

A similar performance in the steady-state behavior of the proposed algorithms at each node $k$ is obtained as shown in Fig. 3. As expected, the DRLP algorithm performs better than all other algorithms included in this comparison.
Fig. 4 shows the MSD at steady-state of the proposed DRLP with those of the DLMP and DMCC under different γ values. We know that the larger γ leads to stronger impulsive noise. One can observe that DRLP are always robust when large outliers occurring (larger γ). This result demonstrates that the proposed DRLP algorithm is much more suitable for dealing with distributed network estimation under different impulsive noise environments.

4 Conclusion

In this paper, we proposed a DRLP algorithm to address the problem of distributed network estimation in the presence of α-stable noise. The DRLP algorithm, based on the minimization of $l_p$-norm, provides an appropriate theoretic to tackle this problem. Furthermore, the complexity of the DRLP is more than one extra nonlinearity computation compared with DRLS. Simulation results show that the DRLP outperforms state-of-the-art solutions for this kind of problems.

References


