Identification of Multilinear Forms Using Combinations of Adaptive Algorithms

Ionuţ-Dorinel Fîciu, Laura-Maria Dogariu, Cristian-Lucian Stanciu, and Constantin Paleologu
University Politehnica of Bucharest, Romania
Email: {ldogariu, cristian, pale}@comm.pub.ro

***Abstract*—The tensor-based adaptive algorithms represent useful tools for the identification of linearly separable systems. These algorithms are designed in the framework of multilinear forms and exploit the decomposition of rank-1 tensors. In this paper, we outline the idea of using different adaptive algorithms (with different performance features) for the individual filters, which represent the components of a rank-1 tensor. In this context, the global scheme based on such combinations could inherit some of the advantages provided by each category of algorithms, e.g., fast convergence rate and low computational complexity.**

***Index Terms*—adaptive filter; multilinear forms; least-mean-square (LMS) algorithm; recursive least-squares (RLS) algorithm; tensor decomposition**

I. Introduction

Many important system identification problems can be efficiently solved using adaptive filters [1]. The real-time features of these signal processing tools are advantageous especially in time-varying environments. In this context, the performance criteria mainly target fast convergence rate, accurate estimation, and low computational complexity.

A more challenging scenario appears in the framework of Multiple-Input Single-Output (MISO) systems. The adaptive filters developed for this purpose should cope with the existence of a large parameter space. However, some of these problems can be formulated in terms of linearly separable systems. Such an approach can be exploited in different applications, like array beamforming, nonlinear acoustic echo cancellation, channel equalization, and source separation, e.g., see [2] and the references therein.

In this context, the tensor-based adaptive filters [2] represent efficient solutions. These adaptive algorithms rely on the decomposition of rank-1 tensors, while the global solution results using a combination of shorter adaptive filters. As shown in [2], the tensorial approach can be applied using the classical adaptive algorithms, e.g., the Least-Mean-Square (LMS) and the Recursive Least-Squares (RLS).

In this work, we explore the idea of using different adaptation modes for the individual filters, aiming to inherit the advantages of each category of algorithms. In other words, the RLS algorithm is used for its fast convergence rate (paid by a higher computational amount), while the normalized LMS (NLMS) algorithm owns the low-complexity feature.

Following this introduction, Section II briefly presents the multilinear framework, while Section III describes the adaptive algorithms and their combination. An experimental result is provided in Section IV, followed by a conclusion in Section V.

II. Identification of Multilinear Forms

In the framework of a real-valued MISO system, the output signal (at discrete-time index n) is defined as

$$y(n) = \sum_{l_1=1}^{L_1} \sum_{l_2=1}^{L_2} \cdots \sum_{l_N=1}^{L_N} x_{l_1 l_2 \ldots l_N}(n) h_{1,l_1} h_{2,l_2} \cdots h_{N,l_N}, \tag{1}$$

where $\mathbf{h}_i = [h_{i,1} \ h_{i,2} \ \cdots \ h_{i,L_i}]^T$ are N individual channels, each one of length $L_i, \ i = 1, 2, \ldots, N$, and the superscript ^{T} denotes the transpose operator. The input signals can be described in the tensorial form $\boldsymbol{\mathcal{X}}(n) \in \mathbb{R}^{L_1 \times L_2 \times \cdots \times L_N}$, having the elements $(\boldsymbol{\mathcal{X}})_{l_1 l_2 \ldots l_N}(n) = x_{l_1 l_2 \ldots l_N}(n)$. Thus, the output signal from (1) becomes

$$y(n) = \boldsymbol{\mathcal{X}}(n) \times_1 \mathbf{h}_1^T \times_2 \mathbf{h}_2^T \times_3 \cdots \times_N \mathbf{h}_N^T, \tag{2}$$

where $\times_i, \ i = 1, 2, \ldots, N$ denotes the mode-i product. As we can notice, $y(n)$ is a multilinear form, since it is a linear function of each $\mathbf{h}_i, \ i = 1, 2, \ldots, N$, when the other $N-1$ components are fixed.

Let us consider the rank-1 tensor $\boldsymbol{\mathcal{H}} \in \mathbb{R}^{L_1 \times L_2 \times \cdots \times L_N}$, with the elements $(\boldsymbol{\mathcal{H}})_{l_1, l_2, \ldots, l_N} = h_{1,l_1} h_{2,l_2} \cdots h_{N,l_N}$, such that

$$\boldsymbol{\mathcal{H}} = \mathbf{h}_1 \circ \mathbf{h}_2 \circ \cdots \circ \mathbf{h}_N, \tag{3}$$

where $\circ$ denotes the outer product. In addition, we have

$$\text{vec}\left(\boldsymbol{\mathcal{H}}\right) = \mathbf{h}_N \otimes \mathbf{h}_{N-1} \otimes \cdots \otimes \mathbf{h}_1, \tag{4}$$

where vec$(\cdot)$ is the vectorization operation and $\otimes$ denotes the Kronecker product. Hence, we can rewrite (1) as

$$y(n) = \text{vec}^T\left(\boldsymbol{\mathcal{H}}\right) \text{vec}\left[\boldsymbol{\mathcal{X}}(n)\right]. \tag{5}$$

Furthermore, we can denote $\mathbf{x}(n) = \text{vec}\left[\boldsymbol{\mathcal{X}}(n)\right]$ and $\mathbf{g} = \text{vec}\left(\boldsymbol{\mathcal{H}}\right)$. Here, $\mathbf{x}(n)$ is the input vector of length $L_1 L_2 \cdots L_N$ and $\mathbf{g}$ plays the role of a global impulse response (of the same length). Therefore, (1) finally becomes

$$y(n) = \mathbf{g}^T \mathbf{x}(n), \tag{6}$$

while the reference signal usually results as

$$d(n) = \mathbf{g}^T \mathbf{x}(n) + w(n), \tag{7}$$

where $w(n)$ is the measurement noise, which is uncorrelated with the input signals. The main goal is the identification of the global system $\mathbf{g}$. Equivalently, the identification problem can be formulated in terms of recursively estimating the individual components $\mathbf{h}_i, \ i = 1, 2, \ldots, N$.

 ISBN: 978-1-61208-858-7

III. Combinations of Adaptive Algorithms

Let us consider the estimated impulse responses of the channels, $\widehat{\mathbf{h}}_i(n),\ i = 1, 2, \ldots, N$, and the estimated output, $\widehat{y}(n)$, such that the error signal result in N equivalent ways:

$$e(n) = d(n) - \widehat{y}(n) = d(n) - \widehat{\mathbf{h}}_i^T(n-1)\mathbf{x}_{\widehat{\mathbf{h}}_i}(n), \quad (8)$$

for $i = 1, 2, \ldots, N$, where

$$\mathbf{x}_{\widehat{\mathbf{h}}_i}(n) = \Big[\widehat{\mathbf{h}}_N(n-1) \otimes \widehat{\mathbf{h}}_{N-1}(n-1) \otimes \cdots \otimes \widehat{\mathbf{h}}_{i+1}(n-1) \otimes \mathbf{I}_{L_i} \otimes \widehat{\mathbf{h}}_{i-1}(n-1) \otimes \cdots \otimes \widehat{\mathbf{h}}_2(n-1) \otimes \widehat{\mathbf{h}}_1(n-1)\Big]^T \mathbf{x}(n),$$

with $\mathbf{I}_{L_i},\ i = 1, 2, \ldots, N$ denoting the identity matrices of sizes $L_i \times L_i$. Using a multilinear optimization strategy based on the mean-squared error (MSE) criterion, the updates of the N adaptive filters result in

$$\widehat{\mathbf{h}}_i(n) = \widehat{\mathbf{h}}_i(n-1) + \mu_i(n)\mathbf{x}_{\widehat{\mathbf{h}}_i}(n)e(n), \quad (9)$$

where $\mu_i(n),\ i = 1, 2, \ldots, N$ are the step-size parameters, while the estimate of the global filter is obtained as

$$\widehat{\mathbf{g}}(n) = \widehat{\mathbf{h}}_N(n) \otimes \widehat{\mathbf{h}}_{N-1}(n) \otimes \cdots \otimes \widehat{\mathbf{h}}_1(n). \quad (10)$$

In nonstationary environments, it is advantageous to follow the line of the NLMS algorithm. In this context, the step-size parameters of the tensor-based NLMS (NLMS-T) algorithm are obtained as $\mu_i(n) = \alpha_i / \left[\delta_i + \mathbf{x}^T_{\widehat{\mathbf{h}}_i}(n)\mathbf{x}_{\widehat{\mathbf{h}}_i}(n)\right]$, with $i = 1, 2, \ldots, N$, where $0 < \alpha_i \leq 1$ are the normalized step-sizes and $\delta_i > 0$ are the regularization constants.

Alternatively, we can apply the least-squares (LS) error criterion [1] in the context of (7) and (8). Thus, the cost functions can be formulated in N alternative ways, following the optimization procedure of the individual impulse responses, Furthermore, the minimization of these cost functions with respect to $\widehat{\mathbf{h}}_i(n),\ i = 1, 2, \ldots, N$ leads to a set of normal equations, which result in the updates of the individual filters:

$$\widehat{\mathbf{h}}_i(n) = \widehat{\mathbf{h}}_i(n-1) + \mathbf{k}_i(n)e(n),\ i = 1, 2, \ldots, N, \quad (11)$$

where $\mathbf{k}_i(n)$ are the Kalman gain vectors and $e(n)$ is evaluated based on (8). The Kalman gain vectors are

$$\mathbf{k}_i(n) = \frac{\mathbf{R}_i^{-1}(n-1)\mathbf{x}_{\widehat{\mathbf{h}}_i}(n)}{\lambda_i + \mathbf{x}^T_{\widehat{\mathbf{h}}_i}(n)\mathbf{R}_i^{-1}(n-1)\mathbf{x}_{\widehat{\mathbf{h}}_i}(n)}, \quad (12)$$

where λ_i ($i = 1, 2, \ldots, N$) are the individual forgetting factors. Finally, the matrix inversion lemma [1] is used to update the matrices $\mathbf{R}_i^{-1}(n)$, i.e.,

$$\mathbf{R}_i^{-1}(n) = \frac{1}{\lambda_i}\left[\mathbf{I}_{L_i} - \mathbf{k}_i(n)\mathbf{x}^T_{\widehat{\mathbf{h}}_i}(n)\right]\mathbf{R}_i^{-1}(n-1), \quad (13)$$

for $i = 1, 2, \ldots, N$. Summarizing, the tensor-based RLS (RLS-T) algorithm is defined by the relations (11)–(13).

In order to take advantage of the particular features of the algorithms, we propose a combination of adaptive filters that uses the RLS-T for the longest filter, while the rest of them are updated as in the NLMS-T algorithm. In this way, the resulting algorithm (namely RLS-NLMS-T) would inherit the fast convergence features of the RLS-T algorithm, while reducing the overall complexity due to the $N-1$ filters that are based on the NLMS-T algorithm.

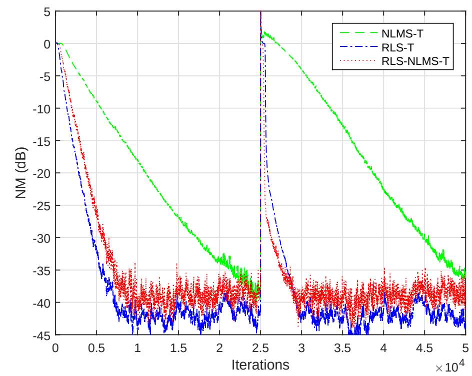

Figure 1. NM of the NLMS-T, RLS-T, and RLS-NLMS-T algorithms.

IV. Experiment

In the following experiment, the input signals are AR(1) processes, which are generated by filtering white Gaussian noises through a first-order system with the pole 0.99. The additive noise $w(n)$ is white and Gaussian, with the variance equal to 0.01. The order of the system used in the experiments is $N = 4$, while the individual impulse responses $\mathbf{h}_i,\ i = 1, 2, \ldots, N$ are generated as in [2], but using $L_1 = 32$, $L_2 = 8$, and $L_3 = L_4 = 4$. Thus, the global impulse response $\mathbf{g}$ has the length 4096. The performance measure is the identification of the global impulse response using the normalized misalignment $\mathrm{NM}\left[\mathbf{g}, \widehat{\mathbf{g}}(n)\right] = \|\mathbf{g} - \widehat{\mathbf{g}}(n)\|_2^2 / \|\mathbf{g}\|_2^2$, where $\|\cdot\|_2$ is the Euclidean norm. In Figure 1, the main parameters of the algorithms are set to $\alpha_i = 0.25$ and $\lambda_i = 1 - 1/50L_i$, for $i = 1, 2, \ldots, N$. As we can notice, the RLS-T and RLS-NLMS-T algorithms perform very similar, in terms of the convergence rate/tracking and misalignment.

V. Conclusion

In this work, we have explored the idea of using a combination of adaptive filters for multilinear forms. The proposed RLS-NLMS-T algorithm achieves a fast convergence rate, while having a lower computational complexity as compared to the RLS-T algorithm. Future works will investigate computationally efficient versions of the RLS-NLMS-T algorithm, which could be based on the coordinate descent iterations [3].

Acknowledgement

This work was supported by two grants of the Romanian Ministry of Education and Research, CNCS–UEFISCDI, project no. PN-III-P1-1.1-PD-2019-0340 and project no. PN-III-P1-1.1-TE-2019-0529, within PNCDI III.

References

[1] S. Haykin, *Adaptive Filter Theory*. Fourth Edition, Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[2] L.-M. Dogariu et al., “Tensor-based adaptive filtering algorithms,” *Symmetry*, vol. 13, id 481 (27 pages), Mar. 2021.

[3] Y. V. Zakharov, G. P. White, and J. Liu, “Low-complexity RLS algorithms using dichotomous coordinate descent iterations,” *IEEE Trans. Signal Processing*, vol. 56, pp. 3150–3161, Jul. 2008.