Signal Processing 98 (2014) 135–142

Contents lists available at ScienceDirect

Signal Processing

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

Variable matrix-type step-size affine projection algorithm with orthogonalized input vectors

PooGyeon Park [a,*], Ji-Hye Seo [b], NamWoong Kong [c]

[a] Department of Electrical Engineering, Pohang University of Science and Technology, Pohang, Gyungbuk 790-784, Republic of Korea
[b] Division of Information Technology of Convergence Engineering, Pohang University of Science and Technology, Pohang, Gyungbuk 790-784, Republic of Korea
[c] POSCO, Pohang, Gyungbuk, Republic of Korea

ARTICLE INFO

Article history:
Received 25 February 2013
Received in revised form 31 October 2013
Accepted 3 November 2013
Available online 18 November 2013

Keywords:
Adaptive filters
Affine projection algorithm
Gram–Schmidt process
Variable matrix-type step size
MSD analysis
Regularization

ABSTRACT

In this paper, we propose a variable matrix-type step-size affine projection algorithm (APA) with orthogonalized input vectors. We generate orthogonalized input vectors using the Gram–Schmidt process to implement the weight update equation of the APA using the sum of normalized least mean squares (NLMS)-like updating equations. This method allows us to use individual step sizes corresponding to each NLMS-like equation, which is equivalent to adopting the step size in the form of a diagonal matrix in the APA. We adopt a variable step-size scheme, in which the individual step sizes are determined to minimize the mean square deviation of the APA in order to achieve the fastest convergence on every iteration. Furthermore, because of the weight vector updated successively only along each innovative one among the reused inputs and effect of the regularization absorbed into the derived step size, the algorithm works well even for badly excited input signals. Experimental results show that our proposed algorithm has almost optimal performance in terms of convergence rate and steady-state estimation error, and these results are remarkable especially for badly excited input signals.

1. Introduction

Adaptive filtering has become an important technique in many applications such as system identification, channel equalization, echo cancellation, and active noise control in the recent decades [1,2]. Least mean squares (LMS) and normalized LMS (NLMS) are the widely used adaptive filtering algorithms owing to their simplicity and robustness. However, one major drawback is that their convergence rate depends on the statistics of their input signal. When a highly correlated signal is applied to an LMS or an NLMS algorithm, it tends to reduce the convergence rate.

One method that can overcome this problem is the affine projection algorithm (APA) [3] suggested by Ozeki and Umeda. This method decorrelates the input signal with a decorrelation degree that depends on the number of input vectors, K, used in the weight update equation. The algorithm speeds up the initial convergence rate, but because it uses multiple input vectors, large steady-state estimation error and high computational complexity problems occur, in contrast to the NLMS algorithm.

In relation to the performance of the APA (in terms of convergence rate and the steady-state estimation error), the step size is a very important factor. A compromise between a fast convergence rate and a small steady-state estimation error can be established when a fixed step size is determined [4]. Thus, the conflict between the convergence rate and steady-state estimation error can be averted by varying the step size. Recently, various

* Corresponding author. Tel.: +82 54 279 2238; fax: +82 54 279 2903.
E-mail addresses: ppg@postech.ac.kr (P. Park), bmclubhs@postech.ac.kr (J.-H. Seo), pandakon@postech.ac.kr (N. Kong).

0165-1684/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2013.11.003

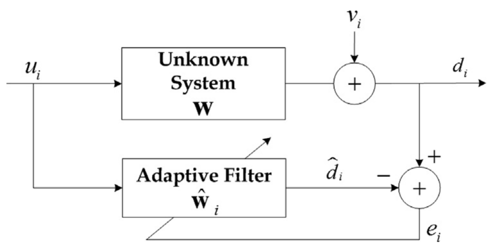

Fig. 1. Adaptive filter structure for system identification.

approaches to varying the step size have been proposed [5–8]. Many of these approaches use algorithms that focus on varying the scalar-type step size at each iteration, resulting in performances that are better than that of fixed step-size APA. Unlike previous algorithms, Paleologu et al. utilized a variable matrix-type step size in the APA [8]. The algorithm extends the concept of step size to a matrix form and makes it possible to consider multiple step sizes at each instant. The method is useful in that the algorithm not only gives better performance than a fixed step-size APA but also involves a novel approach to determining the step size. In the class of the APA, one thing to be considered is a regularization that is necessary for the badly excited input signals to avoid ill-conditioned matrix inversion [1,2]. The algorithm in [8] also adopts a regularization factor for practical use, but it was not included in the derivation of the step size, which leads to some deviation from real-world conditions. Thus, a proper value of regularization factor is determined manually according to various environment, and it leaves us some issues such as deriving more optimal step sizes dealing with behavior of the algorithm for badly excited input signals.

From this motivation, we propose a novel APA that uses a diagonal matrix as its step size to improve the performance of the filter regardless of the excitation level of input signals. The APA can be implemented equivalently with orthogonalized input vectors using Gram–Schmidt orthogonalization [9,10]. It shows the same performance as the conventional APA. However, it structurally modifies the weight update equation because the autocorrelation matrix of the input signal becomes diagonal. Thus, the weight update equation is expressed by the sum of the NLMS-like updating equations. When the APA employs a scalar-type step size, these NLMS-like updating equations have a common step size. However, the structure of the NLMS-like updating equations allows us to use the individual step sizes corresponding to each NLMS-like updating equation. This method is equivalent to expressing the step size in the form of a diagonal matrix. To improve the performance of the filter, the individual step sizes are derived theoretically by analyzing the mean square deviation (MSD) of the APA. Derivation of the step sizes induces the fastest convergence on every iteration, and therefore, we can achieve fast initial convergence and small steady-state estimation error. Furthermore, since we update the weight vector along each innovative one among the reused inputs and derived step sizes contain regularization effects, the proposed algorithm maintains its improved performance even for badly excited input signals.

The rest of this paper is organized as follows. In Section 2, we introduce the structure of the adaptive filter and briefly review the APA implementation with orthogonalized input vectors. Our proposed algorithm is derived in Section 3. In Section 4, several experimental results that indicate that our proposed algorithm outperforms other existing algorithms are discussed. Finally, Section 5 presents our conclusions.

2. Affine projection algorithm with orthogonalized input vectors

Fig. 1 shows the structure of the adaptive filter for system identification. The objective of the adaptive filter is to estimate an n-dimensional weight vector $\mathbf{w}$, which can be obtained by minimizing the error signal e_i, i.e., the difference between the unknown system output d_i and the estimated filter output $\hat{d}_i$ such that $e_i = d_i - \hat{d}_i$. The desired signal d_i and the filter output can be represented as

$$d_i = \mathbf{u}_i^T \mathbf{w} + v_i, \tag{1}$$

$$\hat{d}_i = \mathbf{u}_i^T \hat{\mathbf{w}}_i, \tag{2}$$

where $\mathbf{u}_i = [u_i, u_{i-1}, \ldots, u_{i-n+1}]^T$ denotes an $n \times 1$ column input vector, $\hat{\mathbf{w}}_i$ is an estimate of $\mathbf{w}$ at iteration i, and v_i accounts for the measurement noise. The APA computes $\hat{\mathbf{w}}_{i+1}$ as

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \mathbf{U}_i [\mathbf{U}_i^T \mathbf{U}_i]^{-1} \mathbf{e}_i \tag{3}$$

where

$$\mathbf{U}_i = [\mathbf{u}_i, \mathbf{u}_{i-1}, \ldots, \mathbf{u}_{i-K+1}],$$
$$\mathbf{e}_i = \mathbf{d}_i - \mathbf{U}_i^T \hat{\mathbf{w}}_i,$$
$$\mathbf{d}_i = [d_i, d_{i-1}, \ldots, d_{i-K+1}]^T,$$

and μ is the step size.

Consider a Gram–Schmidt orthogonalization of the reference input vectors $\mathbf{u}_{i-k}$ $(k = 0, 1, \ldots, K-1)$ [11]. Assume that the input vectors $\mathbf{u}_{i-k}$ $(k = 0, 1, \ldots, K-1)$ are independent. Then, the output vectors of the orthogonalization are denoted by $\tilde{\mathbf{u}}_{i,k}$ $(k = 0, 1, \ldots, K-1)$ and they can be generated as follows:

$$\tilde{\mathbf{u}}_{i,k} = \mathbf{u}_{i-k} - \sum_{l=0}^{k-1} \frac{\mathbf{u}_{i-k} T \tilde{\mathbf{u}}_{i,l}}{\|\tilde{\mathbf{u}}_{i,l}\|^2} \tilde{\mathbf{u}}_{i,l}. \tag{4}$$

By rearranging Eq. (4) for $\tilde{\mathbf{u}}_{i,k}$, we get

$$\mathbf{u}_{i-k} = \tilde{\mathbf{u}}_{i,k} + \sum_{l=0}^{k-1} \frac{\mathbf{u}_{i-k} T \tilde{\mathbf{u}}_{i,l}}{\|\tilde{\mathbf{u}}_{i,l}\|^2} \tilde{\mathbf{u}}_{i,l}, \tag{5}$$

and thus the input data matrix $\mathbf{U}_i$ can be expressed as

$$\begin{aligned}\mathbf{U}_i &= \left[\mathbf{u}_i, \mathbf{u}_{i-1}, \ldots, \mathbf{u}_{i-K+1}\right] \\ &= \left[\tilde{\mathbf{u}}_{i,0}, \tilde{\mathbf{u}}_{i,1}, \ldots, \tilde{\mathbf{u}}_{i,K-1}\right] \begin{bmatrix} 1 & \frac{\mathbf{u}_{i-1}^T \tilde{\mathbf{u}}_{i,0}}{\|\tilde{\mathbf{u}}_{i,0}\|^2} & \cdots & \frac{\mathbf{u}_{i-K+1}^T \tilde{\mathbf{u}}_{i,0}}{\|\tilde{\mathbf{u}}_{i,0}\|^2} \\ 0 & 1 & \cdots & \frac{\mathbf{u}_{i-K+1}^T \tilde{\mathbf{u}}_{i,1}}{\|\tilde{\mathbf{u}}_{i,1}\|^2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \\ &= \tilde{\mathbf{U}}_i \cdot \mathbf{L}_i^T, \end{aligned} \tag{6}$$