Signal Processing 115 (2015) 151–156

Contents lists available at ScienceDirect

Signal Processing

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

Fast communication

# Variable step-size affine projection sign algorithm using selective input vectors

Seung Hun Kim [a], Jae Jin Jeong [a], Jong Hyun Choi [a], Sang Woo Kim [b,*]

[a] Department of Electrical Engineering, Pohang University of Science and Technology, Pohang 790-784, Korea

[b] Department of Creative IT Excellence Engineering and Future IT Innovation Laboratory, Pohang University of Science and Technology, Pohang 790-784, Korea

ARTICLE INFO

*Article history:*
Received 14 October 2014
Received in revised form
6 April 2015
Accepted 7 April 2015
Available online 16 April 2015

*Keywords:*
Adaptive filter
Affine projection algorithm
Sign algorithm
Projection order
Mean-square deviation

ABSTRACT

Affine projection sign algorithm (APSA) is a useful adaptive filter for a highly correlated input signal in the presence of impulsive noise. In this study, a novel variable step-size APSA is proposed using selective input vectors to achieve both fast convergence rate and low steady-state mean-square deviation (MSD) with low computational cost. The selective input vectors and step size are chosen so as to maximize the theoretical MSD difference derived using Price's theorem. The simulation results show that the proposed algorithm has the fastest convergence rate and lowest steady-state MSD when compared with recent variable step-size APSAs. Moreover, it effectively reduces computational cost.



## 1. Introduction

Numerous adaptive filtering algorithms with different advantages and disadvantages depending on the design purpose and system environment have been proposed [1]. Among them, the least mean-square (LMS) algorithm and its normalized version are preferred because of their structural simplicity and low computational cost. However, in specific case such as double-talk situation when the input signal is highly correlated and impulsive noise is present, LMS algorithm exhibits degraded performance.

Affine projection algorithm (APA), which uses multiple input vectors to update the weight vector, was introduced to improve the filter performance when the input signal is highly correlated [2]. Additionally, sign algorithm (SA) which was derived from $\mathcal{L}_1$ −norm minimization of error, was employed to design filters robust to impulsive noise [3]. Subsequently, the affine projection sign algorithm (APSA) was introduced for both highly correlated input signal and impulsive noise conditions [4]. As with other AP-type filters, in APSA, the use of multiple input vectors results in fast convergence rate at a high computational cost and memory requirement.

To overcome this, numerous AP-based algorithms including pseudo-type algorithms [5,6], fast-type algorithms [7,8] and selective input vector algorithms [9,11,10] were introduced. Although the pseudo-type and the fast-type algorithms effectively decreased the computational cost through numerical methods, these algorithms are not directly related to the filter's performance such as the convergence rate and the steady-state error. In contrast, the proper choice of input vectors has a relatively weak effect on reducing computational cost, but can improve the filter's performance [9–11]. In APA, there have been active studies about input vector selection algorithms derived from theoretical mean-square error (MSE) or deviation (MSD). However, such APSAs have not yet been introduced, and therefore it is important to derive a new criterion for APSA to give physical meaning for input vectors.

* Corresponding author.
*E-mail addresses:* pisics@postech.ac.kr (S.H. Kim), jin03jin@postech.ac.kr (J.J. Jeong), y2kscore@postech.ac.kr (J.H. Choi), swkim@postech.ac.kr (S.W. Kim).

http://dx.doi.org/10.1016/j.sigpro.2015.04.005
0165-1684/

In addition, step size, which is a scaler for adaptation term, significantly influences convergence rate and steady-state MSD. To achieve both fast convergence rate and low steady-state MSD, many variable step-size APSAs have been introduced recently [12–14]. These algorithms derive the theoretical MSD under the assumption of noise in the transient state [12], by estimating the noise expectation [13] or using a noise bound [14]. However, the physical meaning for specific input vectors is hard to be derived from the MSDs in [12–14], because those MSDs are based on the expectation of $\mathcal{L}_1-$norm of the output error vector.

In this study, we propose a novel variable step-size APSA with selective input vectors. To derive the input selection strategy and the variable step-size scheme, we calculate the theoretical MSD using Price's theorem [15]. The filter employs only selective input vectors which contribute to increase the theoretical MSD difference, and the difference is reconstructed with the selected vectors. In addition, the optimal step size is obtained to maximize the reconstructed MSD difference. The simulation results show that the proposed algorithm has the fastest convergence rate and lowest steady-state error when compared with other variable step-size algorithms [13,14] for various initial projection orders, and the computational cost is effectively decreased.

## 2. Theoretical MSD in APSA

The APSA uses $\mathcal{L}_1-$norm minimization criterion of error to estimate the system weight vector ($\mathbf{w}^o$) [4]. To estimate $\mathbf{w}^o$, the filter weight vector ($\mathbf{w}_{i+1}$) is updated using the current weight vector ($\mathbf{w}_i$), the input matrix $\mathbf{U}_i = [\mathbf{u}_i\ \mathbf{u}_{i-1}\ \ldots \mathbf{u}_{i-K+1}]$ composed of the input vector $\mathbf{u}_i = [u(i)u(i-1)\ldots u(i-M+1)]^T$, and the error vector defined by $\mathbf{e}_i = \mathbf{d}_i - \mathbf{y}_i = [e_1(i)\ e_2(i)\ \ldots e_K(i)]^T$ where $\mathbf{d}_i$ and $\mathbf{y}_i$ are the system output measurement and the filter output at the $i$th iteration, respectively. Further, $M$ is the tap length of the adaptive filter and the projection order ($K$) denotes the number of input vectors used ($K \le M$, [1]). The ordinary weight vector update equation for the APSA of [4] is

$$\mathbf{w}_{i+1} = \mathbf{w}_i + \mu \frac{\mathbf{U}_i\ \mathrm{sgn}(\mathbf{e}_i)}{\sqrt{\mathrm{sgn}(\mathbf{e}_i^T)\mathbf{U}_i^T\mathbf{U}_i\ \mathrm{sgn}(\mathbf{e}_i)}}, \tag{1}$$

where $\mu$ is the step size and $\mathrm{sgn}(\cdot)$ denotes a vector composed of the signs of the target vector. By subtracting (1) from $\mathbf{w}^o$, we obtain the weight error vector $\tilde{\mathbf{w}}_i = \mathbf{w}^o - \mathbf{w}_i$ as

$$\tilde{\mathbf{w}}_{i+1} = \tilde{\mathbf{w}}_i - \mu \frac{\mathbf{U}_i\ \mathrm{sgn}(\mathbf{e}_i)}{\sqrt{\mathrm{sgn}(\mathbf{e}_i^T)\mathbf{U}_i^T\mathbf{U}_i\ \mathrm{sgn}(\mathbf{e}_i)}}. \tag{2}$$

To obtain the MSD equation, we first take the expectation on the squared norm of (2) as

$$E\left[\tilde{\mathbf{w}}_{i+1}^T\tilde{\mathbf{w}}_{i+1}\right] = E\left[\tilde{\mathbf{w}}_i^T\tilde{\mathbf{w}}_i\right] - \Delta_i, \tag{3}$$

where the difference $\Delta_i$ is given by

$$\Delta_i = 2\mu E\left[\frac{\tilde{\mathbf{w}}_i^T\mathbf{U}_i\ \mathrm{sgn}(\mathbf{e}_i)}{\sqrt{\mathrm{sgn}(\mathbf{e}_i^T)\mathbf{U}_i^T\mathbf{U}_i\ \mathrm{sgn}(\mathbf{e}_i)}}\right] - \mu^2.$$

To go further, the numerator and the denominator of $\Delta_i$ can be approximately divided as

$$\Delta_i \approx 2\mu \frac{E\left[\tilde{\mathbf{w}}_i^T\mathbf{U}_i\ \mathrm{sgn}(\mathbf{e}_i)\right]}{E\left[\sqrt{\mathrm{sgn}(\mathbf{e}_i^T)\mathbf{U}_i^T\mathbf{U}_i\ \mathrm{sgn}(\mathbf{e}_i)}\right]} - \mu^2. \tag{4}$$

This separation of the expectation operator is reasonable when the tap length is sufficiently long [16], because the denominator of (4) is slow varying with the long tap. In addition, we apply an approximation as follows:

$$\begin{aligned} &E\left[\sqrt{\mathrm{sgn}(\mathbf{e}_i^T)\mathbf{U}_i^T\mathbf{U}_i\ \mathrm{sgn}(\mathbf{e}_i)}\right] \\ &= E\left[\sqrt{\sum_{j=i-K+1}^{i}\|\mathbf{u}_j\|^2 + \sum_{\substack{m\neq n \\ i-K+1\le m,n\le i}} \mathrm{sgn}(e_m(i))\ \mathrm{sgn}(e_n(i))\mathbf{u}_m^T\mathbf{u}_n}\right] \\ &\approx E\left[\sqrt{\sum_{j=i-K+1}^{i}\|\mathbf{u}_j\|^2}\right]. \end{aligned} \tag{5}$$

Here, we assume that $\sum_{j=i-K+1}^{i}\|\mathbf{u}_j\|^2$ is dominant on (5), and this assumption is reasonable especially in steady-state, because $e_j(i)$ is a zero-mean Gaussian random variable conditioned on $\mathbf{w}_i$ [16,17]. As in [16], we use the assumption even when transient-state for mathematical tractability. Therefore,

$$\Delta_i \approx 2\mu \frac{E\left[\tilde{\mathbf{w}}_i^T\mathbf{U}_i\mathrm{sgn}(\mathbf{e}_i)\right]}{E\left[\sqrt{\sum_{j=i-K+1}^{i}\|\mathbf{u}_j\|^2}\right]} - \mu^2.$$

The numerator is now rearranged as the summation of scalar random variables to apply Price's theorem [15] such as

$$\tilde{\mathbf{w}}_i^T\mathbf{U}_i\ \mathrm{sgn}(\mathbf{e}_i) = \mathbf{e}_{a,i}^T\ \mathrm{sgn}(\mathbf{e}_i) = \sum_{j=1}^{K} e_{a,j}(i)\ \mathrm{sgn}(e_j(i)), \tag{6}$$

where $\mathbf{e}_{a,i} = \mathbf{U}_i^T\tilde{\mathbf{w}}_i = [e_{a,1}(i)e_{a,2}(i)\ldots e_{a,K}(i)]^T$ is the noise-free error vector at the $i$th iteration. By substituting (6) into $\Delta_i$, $\Delta_i$ becomes

$$\Delta_i = 2\mu \frac{\sum_{j=1}^{K} E\left[e_{a,j}(i)\ \mathrm{sgn}(e_j(i))\right]}{E\left[\sqrt{\sum_{j=i-K+1}^{i}\|\mathbf{u}_j\|^2}\right]} - \mu^2. \tag{7}$$

The numerator in (7) is obtained from the following Lemma.

**Lemma.** *When random variables x and y are scalar zero-mean jointly Gaussian, the following statement holds.*

$$E[x\ \mathrm{sign}(y)] = \frac{E[xy]}{\sigma_y^2}E[y\ \mathrm{sign}(y)]$$

**Proof.** The proof could be driven by defining the input–output mapping function in Price's theorem [15] and its extension [1] as $f(x,y) = x\ \mathrm{sign}(y)$. That is,

$$\frac{\partial E[x\ \mathrm{sign}(y)]}{\partial \rho} = E\left[\frac{d\ \mathrm{sign}(y)}{dy}\right] \quad \text{and} \quad E[y\ \mathrm{sign}(y)] = \sigma_y^2 E\left[\frac{d\ \mathrm{sign}(y)}{dy}\right], \tag{8}$$

where $\rho = E[xy]$, and the second equation in (8) was derived from the integration by parts. By integrating the