Variable Step-size Affine Projection Algorithm Based on Excess Mean Square Error

Chang Hee Lee and PooGyeon Park

Electrical and Computer Engineering Division,
Pohang University of Science and Technology,
San 31, Hyojadong, Namgu, Pohang, Kyungbuk, 790-784, Korea
Tel: 82-54-279-5588; Fax: 82-54-279-2903
{eooe, ppg}@postech.ac.kr

Abstract

This paper proposes a computationally reduced method for the step-size update in the variable step-size affine projection algorithm (VSS-APA). Using the previous steady-state analysis and the estimated excess means square error, updating the step-size can be simplified while the advantage of VSS-APA is maintained.

keywords

Adaptive filters, affine projection algorithm, variable step-size

1 Introduction

Signal processing is one of the most important parts in the industrial applications. A useful operation could be control, data transmission, denoising, prediction, filtering and etc. [1]. It is also applicable to the sensor acquisition and the communication among each unit in robotics.

One of the widely used algorithms in this field is the normalized least mean square (NLMS) because of its simple implementation. However it has poor convergence speed for the colored input signal. To overcome this disadvantage, the affine projection algorithm (APA) was suggested by Ozeki and Umeda [2]. By whitening the input signal, the APA can achieve faster convergence speed than NLMS, while the steady-state error of the APA gets worse than that of the NLMS. Reducing step-sizes as time goes on may make the mean-square deviation (MSD) small, which help overcome the disadvantage. As well as, the previously suggested algorithms such as the variable step-size or the variable regularization algorithms can achieve both a fast convergence speed and low steady-state errors simultaneously [3, 4].

In this paper, we focus on the variable step-size APA (VSS-APA) suggested by Shin et al. [4]. Shin's VSS-APA shows good performance over the wide range of system orders, denoted by n, and the input signals, but the projection of the estimation error into the input regression space requires a large amount of computation and the resulting projection vector also has a big dimension (n by 1) that is used in the step-size adjustment at each iteration. As the system order increases, this issue becomes serious. To moderate this problem, this paper suggests an algorithm that estimates the excess mean square error (EMSE) instead of the projection vector. With a scalar value EMSE, a lot of computation is reduced including the projection operation.

To update the instant value of the step-sizes from the EMSE, we use the previous steady-state analysis of the APA. It is known that the steady-state errors are related to the step-size and the variance of the measurement noise in analysis. Moreover, when the step-size is large, it is also affected by the projection order. Since the expected EMSE can be calculated in the analysis for each step-size, the preferable step-size can be inversely obtained from the EMSE. Finally, the required computation gets remarkably reduced while the performance is still quite comparable for any possible input signals. The simulation results verify the performance of the proposed algorithm for various input signals and several system and projection orders.

2 Preliminary

In the APA to be analyzed, the input regression matrix and desired signal sequence is denoted by $\mathbf{U}_i$ and d_i, respectively at time $i \geq 0$. The recursion equa-

©ISAROB 2010

tion of estimated weight $\hat{\mathbf{w}}_i \in \mathcal{R}^{n\times 1}$ for an unknown weight vector $\mathbf{w}_o \in \mathcal{R}^{n\times 1}$ with the system order n and the projection order M is

$$\hat{\mathbf{w}}_{i+1} = \hat{\mathbf{w}}_i + \mu \mathbf{U}_i \left(\epsilon \mathbf{I} + \mathbf{U}_i^T \mathbf{U}_i\right)^{-1} \mathbf{e}_i, \qquad (1)$$

where

$$\mathbf{U}_i = \begin{bmatrix}\mathbf{u}_i & \mathbf{u}_{i-1} & \cdots & \mathbf{u}_{i-M+1}\end{bmatrix} \in \mathcal{R}^{n\times M}, \qquad (2)$$

$$\mathbf{u}_i = \begin{bmatrix}x_i & x_{i-1} & \cdots & x_{i-n+1}\end{bmatrix}^T \in \mathcal{R}^{n\times 1}, \qquad (3)$$

$$\mathbf{e}_i = \mathbf{d}_i - \mathbf{U}_i^T \hat{\mathbf{w}}_i \in \mathcal{R}^{M\times 1} \qquad (4)$$

$$\mathbf{d}_i = \begin{bmatrix}d_i & d_{i-1} & \cdots & d_{i-M+1}\end{bmatrix}^T \in \mathcal{R}^{M\times 1}, \qquad (5)$$

ϵ is the regularization factor, and μ is the step-size. The desired vector from the linear model:

$$\mathbf{d}_i = \mathbf{U}_i^T \mathbf{w}_o + \mathbf{v}_i, \qquad (6)$$

where $\mathbf{v}_i = \begin{bmatrix}v_i & v_{i-1} & \cdots & v_{i-M+1}\end{bmatrix}^T \in \mathcal{R}^{M\times 1}$ is additive measurement noise. Then, the error vector $\mathbf{e}_i$ rewritten as

$$\mathbf{e}_i = \mathbf{d}_i - \mathbf{U}_i^T \hat{\mathbf{w}}_i = \mathbf{U}_i^T \tilde{\mathbf{w}}_i + \mathbf{v}_i, \qquad (7)$$

where $\tilde{\mathbf{w}}_i = \hat{\mathbf{w}}_i - \mathbf{w}_o$. The *a priori* error vector is denoted by $e_{a,i} = \mathbf{u}_i^T \hat{\mathbf{w}}_i$. The EMSE is defined by

$$\text{EMSE} \triangleq \lim_{i\to\infty} \mathrm{E}\left|e_{a,i}\right|^2. \qquad (8)$$

Since a small value of ϵ is assumed in this paper, we start from (22) in [5],

$$\text{EMSE} = \frac{\mu\sigma_v^2}{2-\mu} \frac{\mathrm{Tr}\left(E[\mathbf{A}_i]\right)}{\mathrm{Tr}\left(\mathbf{S}\cdot E[\mathbf{A}_i]\right)}, \qquad (9)$$

where

$$\mathbf{A}_i = (\epsilon\mathbf{I} + \mathbf{U}_i^T\mathbf{U}_i)^{-1}\mathbf{U}_i^T\mathbf{U}_i(\epsilon\mathbf{I} + \mathbf{U}_i^T\mathbf{U}_i)^{-1} \qquad (10)$$

and σ_v^2 is the variance of v_i. When μ is small, identity matrix $\mathbf{I} \in \mathcal{R}^{M\times M}$ can replace $\mathbf{S}$ in (9). Then

$$\text{EMSE} \simeq \frac{\mu\sigma_v^2}{2-\mu}. \qquad (11)$$

When μ is large, $\mathbf{S} \simeq \mathbf{1}\cdot\mathbf{1}^T$ is substituted into (9). Since $\mathrm{Tr}\left[\mathbf{R}_u\right] = \mathrm{Tr}\left[\mathrm{E}\left\{\mathbf{u}_i\mathbf{u}_i^T\right\}\right] = \mathrm{E}\{\|\mathbf{u}_i\|^2\}$, in this case,

$$\text{EMSE} \simeq \frac{\mu\sigma_v^2}{2-\mu}\mathrm{Tr}\left(\mathbf{R}_u\right)\mathrm{E}\left\{\frac{M}{\|\mathbf{u}\|^2}\right\} \qquad (12)$$

$$\simeq \frac{\mu\sigma_v^2 M\|\mathbf{u}_i\|^2}{2-\mu}\mathrm{E}\left\{\frac{1}{\|\mathbf{u}\|^2}\right\}. \qquad (13)$$

When $n \gg 1$, $\mathrm{E}\{1/\|\mathbf{u}_i\|^2\} \simeq 1/\mathrm{E}\{\|\mathbf{u}_i\|^2\}$, which results in

$$\text{EMSE} \simeq \frac{\mu\sigma_v^2 M}{2-\mu}. \qquad (14)$$

3 Variable Step-size

In the previous section, theoretically EMSE is determined for both small and large step-sizes in the case of small ϵ. In this section, EMSE is estimated from the measured error. If v_i is assumed to be independent, identically distributed (i.i.d) and zero-mean, it is obtained that

$$\mathrm{E}\{e_i\} = \mathrm{E}\{e_{a,i} + v_i\} = \mathrm{E}\{e_{a,i}\}. \qquad (15)$$

The following time average ε_i of e_i is calculated by

$$\varepsilon_{i+1} = \alpha\varepsilon_i + (1+\alpha)e_i, \qquad (16)$$

with a smoothing factor $\alpha = 1 - 1/(K)$, where K is a positive integer. When K is close to 1, $\varepsilon_i \simeq e_i$ and When $K \gg 1$, ε_i is close to $\mathrm{E}\{e_{a,i}\}$. In the both cases,

$$\text{EMSE} = \mathrm{E}\left|e_{a,i}\right|^2 \neq \left|\varepsilon_i\right|^2. \qquad (17)$$

However, if we properly select the value of K so that $\varepsilon_i \simeq e_{a,i}$, it is reasonable that

$$\text{EMSE} \simeq |\varepsilon_i|^2. \qquad (18)$$

Using the previous results (11) and (14), the following equation is possible:

$$|\varepsilon_i|^2 \simeq \frac{\mu\beta}{2-\mu}, \qquad (19)$$

where

$$\beta = \begin{cases} \sigma_v^2 & \text{for small } \mu, \quad (20a) \\ M\sigma_v^2 & \text{for large } \mu. \quad (20b) \end{cases}$$

From (19), the step-size update equation is inversely obtained as follows:

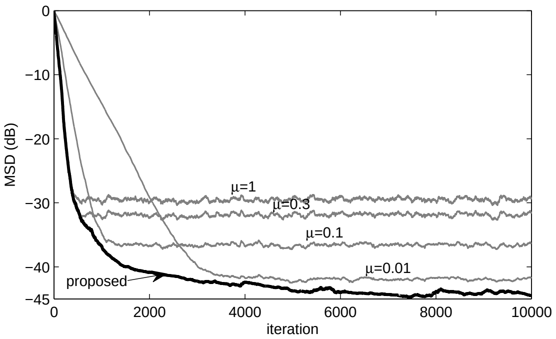
$$\mu_i = \frac{2\left|\varepsilon_i\right|^2}{\beta + \left|\varepsilon_i\right|^2}. \qquad (21)$$

Adjusting β in the run-time, depending on the state of the algorithm, is another issue for the performance improvement, which is out of the scope. Instead of varying β, the constant values are used in the following section. For the performance reason, upper bound 1 in μ is applied [6].

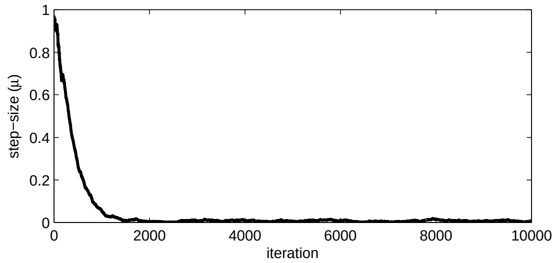
4 Simulations

We conduct several simulations to verify the proposed VSS-APA in a channel estimation scenario. The adaptive filters and the unknown channels are assumed

©ISAROB 2010



(a)



(b)

Figure 1: (a) MSD learning curves of the proposed VSS-APA and the standard APA with different step-sizes between 0.01 and 1 for a zero-mean white Gaussian input signal at $n = 64, M = 4$, SNR = 30 dB $\beta = \sigma_v^2$, and $K = 8$. (b) the corresponding variable step-size of the proposed VSS-APA.

to have the same length of taps in stationary environments. It is defined that $\mathrm{SNR} = 10\log_{10}(\sigma_y^2/\sigma_v^2)$, where σ_y^2 is the variance of output signals. In this section, $n = 64$, $M = 4$, and SNR = 30 dB. The plots are the results of the ensemble average over 50 independent trials.

The first simulation is for verifying how closely the proposed algorithm follows the minimum route of the MSD of the each step-size. The input signal is a zero-mean white Gaussian with a unit variance. As shown in Fig. 1, the MSD learning curve of the proposed VSS-APA has fastest convergence rate that is almost equal to that of the standard APA with $\mu = 1$. It also shows the smaller steady-state error than that of the standard APA with $\mu = 0.01$. Fig. 1b shows how the step-size is varied according to the iteration. At the beginning, large step-sizes is used and after 2000 in the iteration, almost same small step-sizes are used.

The second simulation shows how the performance is affected by K. When $K = 1$, the proposed VSS-APA shows a poor performance in the convergence rate and the steady-state error. As K increases, the convergence rate of the proposed VSS-APA get slower while the steady-state error get smaller, which mean the value K plays an important role in the performance of the proposed algorithm.

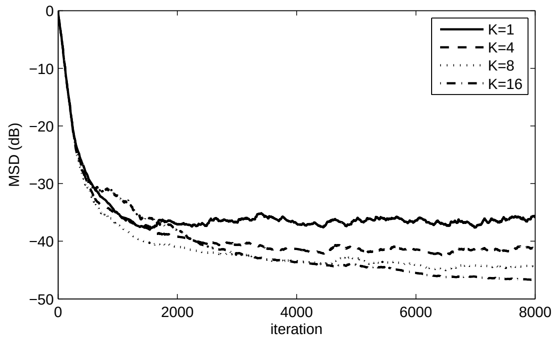


Figure 2: MSD learning curves of the proposed VSS-APA with different values of K at $n = 64$, $M = 4$, SNR = 30 dB, and $\beta = \sigma_v^2$.

The third simulation shows how the performance is affected by β. When $\beta = \sigma_v^2$, the proposed VSS-APA shows apparently fastest convergence rate in Fig. 3 at the cost of larger steady-state error than those of others. It is verified that decreasing the value of β leads smaller steady-state errors.

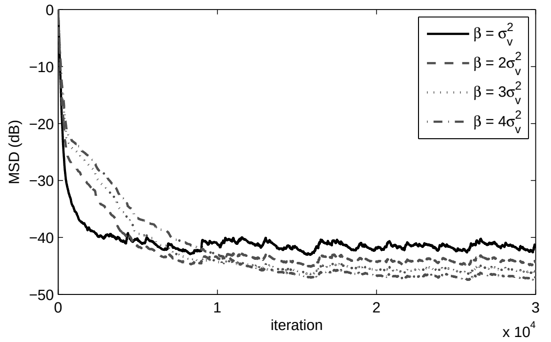


Figure 3: MSD learning curves of the proposed VSS-APA with different values of β at $n = 64$, $M = 4$, SNR = 30 dB, and $K = 4$.

The last two simulations are for performance comparisons to the previous results in the literature [3, 4]. Their tunning parameters are $\alpha = 0.9922$, $C = 1.875 \times 10^{-4}$ in Shin's VSS-APA and $\eta = 1$, $\delta = 0.06$ in Rey's VR-APA. Fig. 4 shows the MSD learning curves of the proposed VSS-APA, Shin's VSS-APA, and Rey's VR-APA when a zero-mean white Gaussian input signal excites the system. The proposed VSS-APA shows the comparable result.

Fig. 5 shows the MSD learning curves of the algorithms when an AR input signal excite the sys-

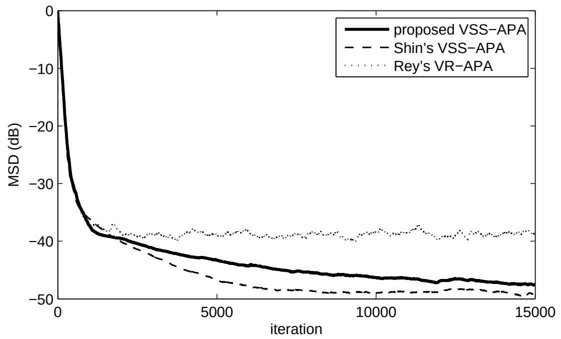


Figure 4: MSD learning curves of the proposed VSS-APA, Shin's VSS-APA and Rey's VR-APA excited by a zero-mean white Gaussian signal at $n = 64$, $M = 4$, SNR = 30 dB, $\beta = 2\sigma_v^2$, and $K = 8$.

tem. The AR input signal is generated by filtering a zero-mean white Gaussian signal through $G_1(z) = 1/(1 - 0.9z^{-1})$. In this case, the performance of the proposed VSS-APA is quite comparable to those of the others.

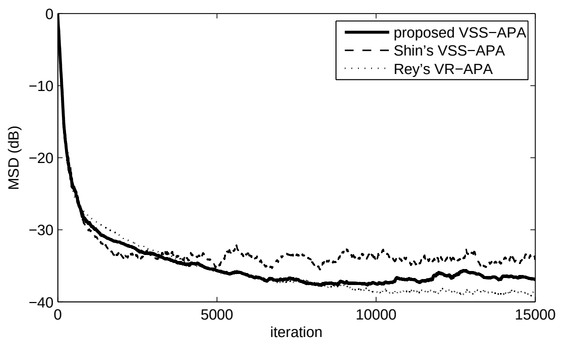


Figure 5: MSD learning curves of the proposed VSS-APA, Shin's VSS-APA and Rey's VR-APA excited by an AR signal at $n = 64$, $M = 4$ SNR = 30 dB, $\beta = \sigma_v^2$, and $K = 8$.

5 Conclusion

This paper proposed a simple way of updating the step-size using the EMSE. The EMSE was estimated from the measurement error and the step-size was inversely derived based on the previous steady-state analysis. As a result, the calculation needed for step-size update became significantly reduced compared to the previous VSS-APA, maintaining its advantage. The simulation results verified the performance of the proposed VSS-APA. It was also shown that additional parameter β led the trade-off between the convergence rate and the steady-state error in the performance.

Acknowledgements

This research was supported by WCU (World Class University) program through the Korea Science and Engineering Foundation funded by the Ministry of Education, Science and Technology (Project No. R31-2008-000-10100-0).

This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency (N I PA-2009-C1090-0902-0004).

This research was supported by the Brain Korea 21 Project in 2009.

References

[1] Haykin, S. (2001) Adaptive Filter Theory, Prentice Hall, NJ 4th edition.

[2] Ozeki, K. and Umeda, T. (1984) *Electronics and communications in Japan* **67(5)**, 19–27.

[3] Rey, H., Vega, L. R., Tressens, S., and Benesty, J. (2007) *IEEE Transactions on Signal Processing* **55(5)**, 2096–2109.

[4] Shin, H. C., Sayed, A. H., and Song, W. J. (2004) *IEEE Signal Processing Letters* **11(2)**, 132–135.

[5] Shin, H. C. and Sayed, A. H. (2004) *IEEE Transactions on Signal Processing* **52(1)**, 90–102.

[6] Sankaran, S. G. and Beex, A. A. (2000) *IEEE Transactions on Signal Processing* **48(4)**, 1086–1096.