

CONVERGENCE ANALYSIS OF SIGN-SIGN LMS ALGORITHM FOR ADAPTIVE FILTERS WITH CORRELATED GAUSSIAN DATA

Byung-Eul Jun, Dong-Jo Park and Yong-Woon Kim

Department of Electrical Engineering
Korea Advanced Institute of Science and Technology (KAIST)
373-1 Kusong-dong Yusong-gu Taejeon, 305-701, Korea

ABSTRACT

This paper presents a statistical behavior analysis of a sign-sign least mean square algorithm, which is obtained by clipping both the reference input signal and the estimation error, for adaptive filters with correlated Gaussian data. The study focuses on derivation of expressions for the first and second moment behaviors of the filter coefficient vector and analysis for the mean square error of the filter. The previous analysis of this type for the sign-sign algorithm is based on the assumption that the input sequence to the adaptive filter is independent, identically distributed Gaussian, but this restriction is removed in our analysis. Theoretical expressions derived are verified numerically through computer simulations for an example of system identification.

1. INTRODUCTION

The least mean square(LMS) adaptive filter algorithm [1] is very popular for its simplicity, but even simpler approaches are required for many real time applications. Replacing the input regressor vector and the estimation error components of the update term by their signs reduces computing time and dynamic range requirements by turning multiplications into bit shifts. Such a variant of the LMS adaptive filter is known as the sign-sign algorithm(SSA) which appears early in [2] as a suggestion of the applicability for use in channel equalization. The sign-sign algorithm has seen a resurgence of interest since its incorporation in a C-CITT standard [3] for adaptive differential pulse code modulation(ADPCM). There are some analysis results on the sign-sign algorithm [4], [5], [6], [7], [8], and the results except [6] are focused on the stability and the persistent excitation condition for the algorithm.

In this paper we analyze the expected behavior of the sign-sign algorithm for the finite impulse response adaptive filters with correlated Gaussian data. This

paper has two main results. The first is the derivation and analysis of the dynamic equations describing the mean and the mean square behavior of the filter coefficients. As the second contribution, we derive the expressions for the mean square error of the filter and analyze the equations. From the analysis we find the relationship between the step size and the mean square error, which gives us a guideline for design of adaptive filters. The results, which are derived from the analysis, are verified numerically through computer simulations for an example of adaptive system identification.

Even though the expected behavior analysis of this type has been active for the LMS [9], [10], the signed-error algorithm [11] and the signed-regressor algorithm [12], there is only one result [6] on the sign-sign algorithm as far as we know. The analysis [6] assumed that the reference input should be independent, identically distributed(i.i.d.). But the i.i.d. condition for the input data is removed in our analysis.

2. CONVERGENCE ANALYSIS

The sign-sign LMS algorithm [2] is given by

$$H(n+1) = H(n) + \mu \operatorname{sgn}\{X(n)\} \operatorname{sgn}\{e(n)\}, \quad (1)$$

$$e(n) = d(n) - H(n)^T X(n), \quad (2)$$

where $H(n) \in \mathcal{R}^N$ is the filter coefficient vector at time n , $X(n) \in \mathcal{R}^N$ is the regressor vector composed of the reference input samples $x(n)$, $d(n)$ is the primary input signal or the desired response of the filter, $(\cdot)^T$ means the transpose of (\cdot) , $e(n)$ is the estimation error, μ is the step size, and $\operatorname{sgn}\{\cdot\}$ means the signum function. The desired response $d(n)$ can be decomposed into the information correlated with the reference input $X(n)$ and the uncorrelated noise as

$$d(n) = H_{opt}^T X(n) + e_{min}(n), \quad (3)$$

where $H_{opt} \in \mathcal{R}^N$ is the optimal coefficient vector or the Wiener solution, and $e_{min}(n)$ is the uncorrelated

noise or the residual error of the Wiener filter. We assume that $e_{min}(n)$ is a Gaussian noise with zero mean and finite variance. Using (3) to (2) produces an expression for the estimation error written by

$$e(n) = e_{min}(n) - V^T(n)X(n), \quad (4)$$

where $V(n) \in \mathcal{R}^N$ is the coefficient error vector defined by $V(n) = H(n) - H_{opt}$

For the convergence analysis, we will assume that $d(n)$ and $x(n)$ are jointly Gaussian, zero mean, stationary signals. As in many convergence analyses of this type [9], [10], [12], [11], we will also assume that the input pairs $\{X(n), d(n)\}$ are mutually uncorrelated for different values of n , which is known as the independence assumption. Note that the assumption does not restrict the nature of the autocorrelation matrix of the input regressor vector $X(n)$.

We can rewrite the sign-sign algorithm (1) by using the filter coefficient error vector $V(n)$ as follows:

$$V(n+1) = V(n) + \mu \operatorname{sgn}\{X(n)\} \operatorname{sgn}\{e(n)\}. \quad (5)$$

Expectation of both sides of (5) gives us

$$E[V(n+1)] = E[V(n)] + \underbrace{\mu E[\operatorname{sgn}\{X(n)\} \operatorname{sgn}\{e(n)\}]}_{\mathbf{A}}, \quad (6)$$

where $E[\cdot]$ is an expectation operator. Under the independence assumption we can write \mathbf{A} as

$$\mathbf{A} = E[\underbrace{E[\operatorname{sgn}\{X(n)\} \operatorname{sgn}\{e(n)\} | V(n)]}_{\mathbf{B}}]. \quad (7)$$

The i th component of \mathbf{B} can be evaluated based on the Price theorem [13] as

$$\mathbf{B}_i = \frac{2}{\pi} \sin^{-1}(\delta_i), \quad (8)$$

where δ_i is a correlation coefficient defined by

$$\delta_i = \frac{E[x(n-i+1)e(n) | V(n)]}{\sigma_x \sigma_{e|V(n)}} = -\frac{\mathbf{r}_i^T V(n)}{\sigma_x \sigma_{e|V(n)}}. \quad (9)$$

In (5), $\sin^{-1}(\cdot)$ is the arcsine function, σ_x is a standard deviation of $x(n)$, $\sigma_{e|V(n)}$ is the standard deviation of the estimation error conditioned by the coefficient error vector and $\mathbf{r}_i \in \mathcal{R}^N$ is the i th column vector of the input autocorrelation matrix $R_{xx} = E[X(n)X^T(n)] \in \mathcal{R}^{N \times N}$. The absolute value of δ_i is less than or equal to 1 when the estimation error has finite variance. Now let us consider an approximation of (8) as

$$\mathbf{B}_i \simeq -\frac{2}{\pi} \frac{1}{\sigma_x \sigma_{e|V(n)}} \mathbf{r}_i^T V(n), \quad (10)$$

where $|\delta_i| \ll 1$ is assumed. Even though the assumption is valid strictly when the algorithm converges to the vicinity of the Wiener solution under the small step size condition, it produces a reasonable approximation in the transient phase. We will confirm the validity and usefulness of the approximation indirectly through computer simulations later. Since it is known in [11] that an approximation, $\sigma_{e|V(n)} \simeq \sigma_e(n)$, is valid and useful under the small step size condition, (7) can be expressed as

$$\mathbf{A} \simeq -\frac{2}{\pi} \frac{1}{\sigma_x \sigma_e(n)} R_{xx} E[V(n)]. \quad (11)$$

Inserting (11) into (6) produces an expression for the mean behavior of the filter coefficient error vector as follows:

$$E[V(n+1)] = \left(I - \frac{2}{\pi} \frac{\mu}{\sigma_x \sigma_e(n)} R_{xx} \right) E[V(n)]. \quad (12)$$

Remark 1: When the reference input is i.i.d., the expression (12) is exactly matched with the Duttweiler's result [6] given by

$$E[V(n+1)] = \left(1 - \frac{2}{\pi} \mu \frac{\sigma_x}{\sigma_e(n)} \right) E[V(n)]. \quad (13)$$

□

We can also evaluate a propagation equation for the autocorrelation matrix of $V(n)$ by using (5) and its transpose as follows:

$$K(n+1) = K(n) - \frac{2}{\pi} \frac{\mu}{\sigma_x \sigma_e(n)} \{R_{xx} K(n) + K(n) R_{xx}\} + \frac{2}{\pi} \mu^2 P_{xx}, \quad (14)$$

where $K(n) \in \mathcal{R}^{N \times N}$ is an autocorrelation matrix of $V(n)$ and $P_{xx} \in \mathcal{R}^{N \times N}$ is a symmetric matrix composed of its components

$$P_{xx}(i, j) = \sin^{-1} \left\{ \frac{R_{xx}(i, j)}{\sigma_x^2} \right\}. \quad (15)$$

Remark 2: Under the i.i.d. input condition, Duttweiler [6] derived an expression corresponding to (14) as

$$K(n+1) = \left(1 - \frac{2}{\pi} \mu \frac{\sigma_x}{\sigma_e(n)} \right)^2 K(n) + \mu^2 I. \quad (16)$$

But (16) has a superfluous term compared with that reduced from (14) for the i.i.d. input data. □

An orthonormal matrix $Q \in \mathcal{R}^{N \times N}$, composed of the normalized orthogonal eigenvectors of R_{xx} , is able to diagonalize the matrix R_{xx} as

$$Q^T R_{xx} Q = \Lambda = \operatorname{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}, \quad (17)$$

where λ_i is the i th eigenvalue of R_{xx} and $\text{diag}\{\dots\}$ means the diagonal matrix composed of $\{\dots\}$. By using Q , (14) can be transformed as follows:

$$K'(n+1) = K'(n) - \frac{2}{\pi} \frac{\mu}{\sigma_x \sigma_e(n)} \left\{ \Lambda K'(n) + K'(n) \Lambda \right\} + \frac{2}{\pi} \mu^2 P'_{xx}, \quad (18)$$

where $K'(n)$ and P'_{xx} are defined by $K'(n) = Q^T K(n) Q$ and $P'_{xx} = Q^T P_{xx} Q$, respectively.

From the definition of the estimation error (4), the mean square error (MSE) can be expressed as

$$\xi(n) = \xi_{min} + \text{tr}\{K(n)R_{xx}\}, \quad (19)$$

where ξ_{min} is the minimum MSE due to the uncorrelated noise $e_{min}(n)$ and the second term is the excess MSE which will be written by $\xi_{ex}(n) = \text{tr}\{K(n)R_{xx}\}$ later. A similarity transformation by Q gives us an expression for the excess MSE as

$$\xi_{ex}(n) = \text{tr}\{K'(n)\Lambda\} = \underline{\lambda}^T \underline{k}'(n), \quad (20)$$

where $\underline{\lambda} \in \mathcal{R}^N$ and $\underline{k}'(n) \in \mathcal{R}^N$ are composed of the diagonal elements of Λ and $K'(n)$, respectively.

If $K'(n)$ converges, then so does the mean square error since the stability of (20) is governed by the convergence of $\underline{k}'(n)$. It is interesting to find an expression for the steady-state mean square error. Getting a steady-state form of (18) as $n \rightarrow \infty$ and summing the equations from $i = 1$ to N , we can get

$$\xi_{ex}(\infty) = \frac{1}{2} \mu \sigma_x \sigma_e(\infty) \text{tr}\{P'_{xx}\}. \quad (21)$$

Using the relation $\text{tr}\{P'_{xx}\} = \text{tr}\{P_{xx}\} = \pi N/2$ for (21) produces

$$\xi_{ex}(\infty) = \frac{\pi}{4} \mu \sigma_x \sigma_e(\infty) N. \quad (22)$$

Expression (22) is a second order equation of the excess MSE written by

$$\xi_{ex}^2(\infty) - \beta^2 \xi_{ex}(\infty) - \beta^2 \xi_{min} = 0, \quad (23)$$

where β is defined by

$$\beta = \frac{\pi}{4} \mu \sigma_x N. \quad (24)$$

Therefore, the excess MSE is given by

$$\xi_{ex}(\infty) = \beta \left[\frac{1}{2} \beta + \left\{ \frac{\beta^2}{4} + \xi_{min} \right\}^{1/2} \right]. \quad (25)$$

3. COMPUTER SIMULATIONS

For verifying the expressions derived above, we will employ a system identification example. In this example we choose the eigenvalue spread ratio of the input autocorrelation matrix R_{xx} to be fairly large, approximately 673, so that the independence assumption is seriously violated. We, in this tough condition, will verify the theoretical results derived under the independence assumption.

The reference input $x(n)$ to the unknown system is given by a fourth-order autoregressive signal,

$$x(n) = 1.79x(n-1) - 1.9425x(n-2) + 1.27x(n-3) - 0.5x(n-4) + \zeta(n), \quad (26)$$

where $\zeta(n)$ is a zero mean white Gaussian random variable with variance such that variance of $x(n)$ is 1. The primary input $d(n)$ is generated from the finite impulse response system with the reference input $x(n)$, the coefficients

$$H_{opt} = [0.1 \ 0.3 \ 0.5 \ 0.7 \ 0.9 \ 0.7 \ 0.5 \ 0.3 \ 0.1]^T \quad (27)$$

and an additive noise, a white Gaussian with zero mean and variance 0.01. The adaptive filter has also the number of taps same with that of the optimal weights (27).

The simulated and theoretical results for the mean square error are depicted on Fig. 1, where the theoretical curve is calculated directly from (14), (19) and the empirical curve is Monte Carlo simulation results which are ensemble averages of 100 independent runs using 10000 data samples each. We can see in Fig. 1 that the theoretical equation derived is predicting closely the empirical behaviors even though the independence assumption is seriously violated.

Table I shows the excess mean square error and misadjustment, where the theoretical values are evaluated by (19) and the simulated ones are the time average from 5001th to 10000th data of the empirical curve in Fig. 1.

TABLE I
STEADY-STATE MEAN SQUARE ERROR IN
STATIONARY ENVIRONMENTS OF IDENTIFICATION
EXAMPLE

	Excess MSE $\xi_{ex}(\infty)$	Misadjustment $\xi_{ex}(\infty)/\xi_{min}$
Theory, eq. (25)	1.000e-3	10.00 %
Simulation	1.020e-3	10.20 %

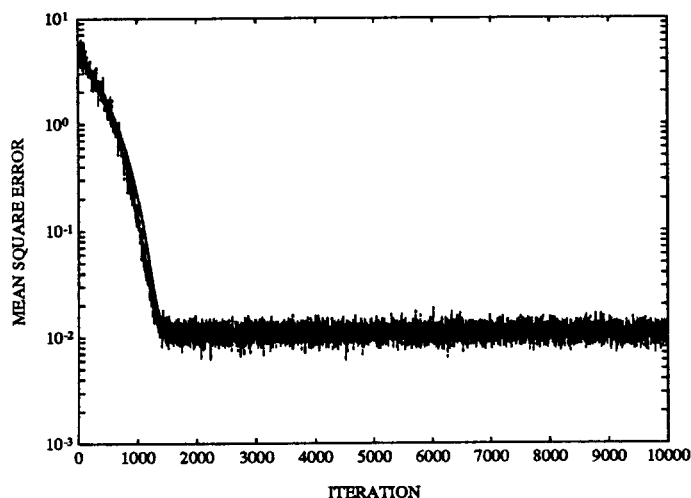


Fig. 1 Mean square error behavior of an identification example: $\xi_{\min} = 0.01$, $\mu = 0.001415$ (for 10 % misadjustment); solid line - theoretical results, dashed line - Monte Carlo simulation results.

4. CONCLUSIONS

We derived and analyzed expressions describing the mean and the mean square behaviors of the sign-sign LMS algorithm for adaptive filters with correlated Gaussian input data.

In this study we get the following expressions and properties on the sign-sign algorithm: The mean behavior of the filter coefficients is expressed as a state equation, where the state vector is defined by the mean of the filter coefficient error vector; For the second moment behavior of the filter coefficients, we derived a propagation equation for the autocorrelation matrix of the coefficient error vector; The expression for the mean square error will give us useful guidelines for design and analysis of adaptive filter systems.

The expressions and properties produced from the analysis were verified numerically through computer simulations for a system identification example, and also the usefulness and validity of the assumptions employed for the analysis were confirmed through the simulations.

5. REFERENCES

- [1] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Englewood Cliffs, N.J., 1985.
- [2] R. W. Lucky, "Techniques for adaptive equalization of digital communication systems," *Bell Syst. Tech. J.*, vol. 45, pp. 225-286, Feb. 1966.
- [3] CCITT Red Book, Recommendation G721, Rome III-3, Oct. 1984.
- [4] S. Dasgupta and C. R. Johnson, Jr., "Some comments on the behavior of sign-sign adaptive identifiers," *Syst. Contr. Lett.*, vol. 7, no. 4, pp. 75-82, Apr. 1986.
- [5] S. Dasgupta, C. R. Johnson, Jr., and A. M. Baksho, "Sign-sign LMS convergence with independent stochastic inputs," *IEEE Trans. Inform. Theory*, vol. 36, no. 1, pp. 197-201, Jan. 1990.
- [6] D. L. Duttweiler, "Adaptive filter performance with nonlinearities in the correlation multilayer," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-30, no. 4, pp. 578-586, Aug. 1982.
- [7] C. R. Elevitch, W. A. Sethares, G. J. Rey, and C. R. Johnson, Jr., "Quiver diagram and signed adaptive filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 2, pp. 227-236, Feb. 1989.
- [8] W. A. Sethares, "The least mean square family," in *Adaptive system identification and signal processing algorithms*, N. Kalouptsidis and S. Theodoridis, Ed. Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [9] A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 1, pp. 222-230, Feb. 1985.
- [10] L. L. Horowitz and K. D. Senne, "Performance advantage of complex LMS for controlling narrow-band adaptive arrays," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, no. 3, pp. 722-736, June 1981.
- [11] V. J. Mathews and S. H. Cho, "Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, no. 4, pp. 450-454, Apr. 1987.
- [12] B. E. Jun, K. S. Eom, and D. J. Park, "Statistical behavior analysis of a signed-regressor LMS algorithm for adaptive filters with correlated Gaussian data," submitted to *IEEE Trans. Signal Processing*, 1994.
- [13] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *IEEE Trans. Inform. Theory*, vol. IT-4, pp. 69-72, June, 1958.