Performance Analysis of Adaptive Filters Using the Sign Algorithm

Elias Masry
Department of Electrical & Computer Engineering, University of California at San Diego, La Jolla, CA 92039

Francesco Bullo
Department of Electrical Engineering, University of Padova, Padova, ITALY

Abstract — Convergence analysis of the sign algorithm for adaptive filtering with Gaussian uncorrelated input processes is presented. Asymptotic time-averaged results for the mean-square deviation error and for the signal estimation error are established. The results hold for arbitrary step size \( \mu > 0 \).

I. INTRODUCTION

Adaptive linear estimation methods based on the principle of steepest descent and its variations have been applied to a wide range of problems such as filtering, noise canceling, line enhancement, antenna processing, and interference suppression. In this paper we are concerned with the convergence analysis of the sign algorithm [1] [3]. The updated equation for the vector \( \mathbf{h}(j) \) of the estimated filter’s coefficients at iteration \( j \) is given by

\[
\mathbf{h}(j+1) = \mathbf{h}(j) + \mu \mathbf{x}(j) \text{sign} \left( \mathbf{e}(j) \right), \quad j = 1, 2, \cdots
\]

(1)

where \( \mathbf{x}(j) \) is the data at iteration \( j \), \( \mathbf{e}(j) \) is the error in estimating the desired signal \( \mathbf{d}(j) \) using the data vector \( \mathbf{x}(j) \),

\[
\mathbf{e}(j) = \mathbf{d}(j) - \mathbf{h}^T(j) \mathbf{x}(j),
\]

(2)

and \( \mu \) is the adaptation size. Here \( \mathbf{h}(j) \) and \( \mathbf{x}(j) \) are column vectors with dimension \( N \); \( \mathbf{h}^T(j) \) is the transpose of \( \mathbf{h}(j) \). The initial estimate \( \mathbf{h}(1) \) is assumed to be nonrandom. The processes \( \{ \mathbf{d}(j) \}_{j=1}^\infty \) and \( \{ \mathbf{x}(j) \}_{j=1}^\infty \) are jointly stationary with finite moments. Let \( R = E(\mathbf{x}(j)\mathbf{x}^T(j)) \) assumed to be positive definite with eigenvalues \( 0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N \). Let \( \mathbf{h}_{opt} \) be the optimal linear Wiener-Hopf filter with corresponding minimum mean-square error \( \varepsilon_{\text{min}}^2 \). Let

\[
\mathbf{e}(j) = \mathbf{h}_{opt} - \mathbf{h}(j)
\]

(3)

be the deviation error in estimating the Wiener-Hopf coefficients. Then the mean-square deviation error for the filter’s coefficients (MSD) is given by \( E[\|\mathbf{e}(j)\|^2] \). \( E[\varepsilon^2(j)] \) is the mean-square error for the signal estimation. We have

\[
E[\varepsilon^2(j)] = \varepsilon_{\text{min}}^2 + \varepsilon_0^2
\]

where \( \varepsilon_0^2(j) \) is the signal estimation "excess mean-square error". Gerstho [1] provided a rigorous analysis of the sign algorithm and showed that

\[
\limsup_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} E[\|\mathbf{e}(j)\|^2] \leq \varepsilon_{\text{min}} + \frac{1}{2\mu} \text{tr}[R]
\]

(4)

for any step size \( \mu > 0 \), where \( \varepsilon_{\text{min}} \) is the least mean-square error minimizing \( E[\|\mathbf{d}(j) - \mathbf{h}^T(j) \mathbf{x}(j)\|^2] \). In Mathews and Cho [3] it is further assumed that \( \{\mathbf{x}(j), \mathbf{d}(j)\}_{j=1}^{\infty} \) are jointly Gaussian i.i.d. random variables with zero means. They provided a heuristic argument, based on an unverifiable approximation, and obtained convergence results for small step size \( \mu \). The purpose of this paper is to establish the asymptotic time-averaged convergence for the mean-square deviation error \( E[\|\mathbf{e}(j)\|^2] \) and for the signal estimation error \( E[\varepsilon^2(j)] \). These results hold for any step size \( \mu > 0 \).

II. RESULTS

We make the assumption that \( \{\mathbf{x}(j), \mathbf{d}(j)\}_{j=1}^{\infty} \) are jointly Gaussian i.i.d. random variables with zero means. The proof of the following results can be found in [2].

\[
\limsup_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} E[\|\mathbf{e}(j)\|^2] \leq C_1 \mu + C_2 \mu^2
\]

(5a)

where the constants \( C_1 \) and \( C_2 \) are given by

\[
C_1 = \frac{9}{2} \sqrt{\frac{\text{tr}[R]}{2}} \frac{\lambda_1}{\lambda_1 - \lambda_2} \epsilon_{\text{min}}
\]

(5b)

\[
C_2 = \frac{27 \pi}{8} \frac{\lambda_N}{\lambda_1} \left[ \frac{\lambda_1 \lambda_N}{\lambda_1 - \lambda_N} \right] \text{tr}[R] + 2 \lambda_N \epsilon_{\text{min}}
\]

(5c)

with \( m_4 = E[\|\mathbf{x}(j)\|^4] \).

\[
\limsup_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} E[\varepsilon^2(j)] \leq C_3 \mu + C_4 \mu^2
\]

(6)

where \( C_3 = \lambda_N C_1 \) and \( C_4 = \lambda_N C_2 \).

REFERENCES

