비선형 무한응답 최소자승형 적응여파기의 수렴속도에 관한 연구

○김 화 종, 은 종 관

한국과학기술원 전기 및 전자공학과

On the Convergence Speed of Nonlinear Least-Squares IIR Adaptive Filter

Hwa Jong Kim and Chong Kwan Un

Department of Electrical Engineering, KAIST

ABSTRACT

In this paper, we investigate an infinite impulse response (IIR) adaptive digital filter (ADF) based on the nonlinear least-squares (NLS) algorithm, and compare its convergence speed to that of a self-orthogonalizing IIR ADF which is known to have fastest convergence. By simulation, it is shown that the NLS IIR ADF converges faster than other known IIR ADF's, especially for a low-order case.

I. INTRODUCTION

Recently, IIR ADF has attracted considerable attention of many researchers because of its several inherent advantages over the finite impulse response (FIR) ADF [1]-[4]. An IIR filter can have a smaller number of coefficients than an FIR filter for the same performance, and it usually matches physical systems well. However, the IIR ADF also has disadvantages, such as the instability problem and the possible local minimization problem with a filter of reduced order [1]. Another disadvantage is that in general it is known to converge slower than the FIR ADF.

In this paper, we study an NLS IIR ADF that can speed up the convergence rate. We also compare its convergence speed to that of the recursive Gauss-Newton (RGN) IIR ADF which is of a self-orthogonalizing type [3].

II. NONLINEAR LEAST-SQUARES IIR ADF

Here we first consider the output error for the IIR ADF shown in Fig.1. Let $d(n)$ be the desired signal defined

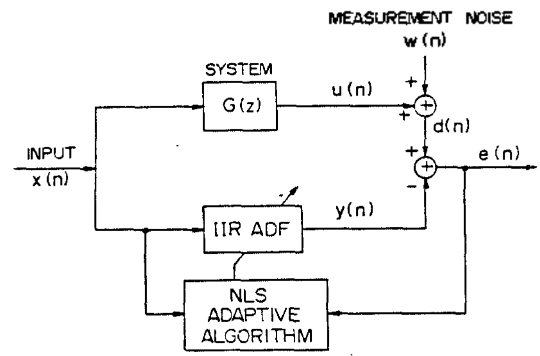

그림 1 시스템 인식을 위한 무한응답 적응 여파기

Fig. 1 Adaptive filter for system identification

by

$$d(n) \triangleq u(n) + w(n) \tag{1}$$

where $u(n)$ is the system output driven by zero-mean white Gaussian noise (WGN) input $x(n)$, and $w(n)$ is the disturbance representing the measurement error which is also an additive WGN. The *a priori* output estimate $y(n)$ of $d(n)$ is then given by

$$\begin{aligned} y(n) &= \sum_{i=1}^{N} a_i(n-1)\bar{y}(n-i) + \sum_{j=0}^{M} b_j(n-1)x(n-j) \\ &= \theta(n-1)^T \phi(n) \end{aligned} \tag{2}$$

where $\{\bar{y}(\cdot)\}$ are *a posteriori* output estimates which will be defined later, and $\{a_i(\cdot)\}$ and $\{b_j(\cdot)\}$ are coefficients of IIR ADF. In (2), the coefficient and information vectors are defined, respectively, by

$$\theta(n-1) \triangleq [a_1(n-1), a_2(n-1), \ldots, a_N(n-1); \; b_0(n-1), \ldots, b_M(n-1)]^T \tag{3}$$

$$\phi(n) \triangleq [\hat{y}(n-1), \hat{y}(n-2), \ldots, \hat{y}(n-N); \; x(n), x(n-1), \ldots, x(n-M)]^T \tag{4}$$

where T denotes transpose. Consequently, the *a priori* output error is given by

$$e(n) = d(n) - y(n) \; . \tag{5}$$

To update the coefficient vector, we use a modified form of the NLS algorithm [4] as the following:

$$\theta(n) = \theta(n-1) + P(n)\psi(n)[d(n)-y(n)] \tag{6}$$

where

$$P(n) = \frac{1}{\lambda}\left[P(n-1) - \frac{P(n-1)\psi(n)\psi(n)^T P(n-1)}{\lambda + \psi(n)^T P(n-1)\psi(n)} \right] \tag{7}$$

and $P(0)$ is an (N+M+1) by (N+M+1) identity matrix. In (7), λ is a weighting factor and the gradient vector $\psi(n)$ is given approximately by

$$\psi(n) \simeq [\hat{y}^f(n-1), \hat{y}^f(n-2), \ldots, \hat{y}^f(n-N); \; x^f(n), x^f(n-1), \ldots, x^f(n-M)]^T \tag{8}$$

where

$$x^f(n) \triangleq x(n) + \sum_{i=1}^{N} a_i(n-1)x^f(n-i) \tag{9}$$

$$\hat{y}^f(n) \triangleq \hat{y}(n) + \sum_{i=1}^{N} a_i(n-1)\hat{y}^f(n-i) \tag{10}$$

$$\hat{y}(n) \triangleq \theta(n)^T\phi(n) \; . \tag{11}$$

It is noted that $\hat{y}(n)$ is *a posteriori* output estimate because it depends on $\theta(n)$. In the next section we will show and discuss simulation results.

III. SIMULATION RESULTS AND DISCUSSION

First, as a low-order case, we investigate a system whose transfer function is given by

$$G(z) = \frac{K}{1-1.2z^{-1}+0.6z^{-2}} \; . \tag{12}$$

In simulation, the constant K is chosen such that the system output has unity power. For measurement noise $w(n)$, we use WGN with power of 0.001. It is assumed that the order of the system is known *a priori* (i.e., N=2 and M=0). In Fig.2, we compare the mean-square *a priori* output errors (MSE's)

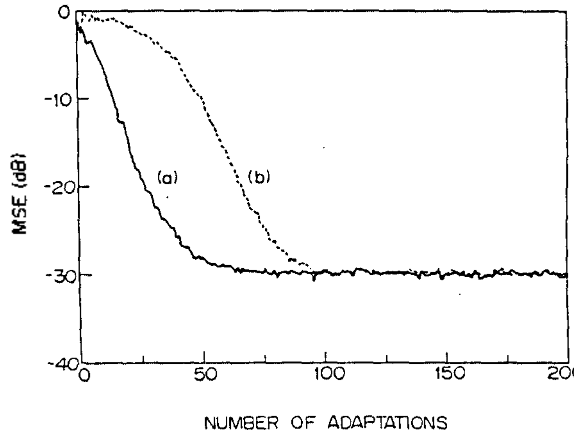

그림 2 극점이 두개인 무한응답 적응여파기의 수렴특성 (앙상블 평균을 위하여 독립적으로 500번 수행하였음)

Fig. 2 MSE of IIR ADF with 2 poles (500 independent iterations were done for an ensemble average)
(a) NLS IIR ADF with λ = 0.95
(b) Modified RGN IIR ADF with α = 0.05

of the proposed NLS IIR ADF and the modified RGN IIR ADF[1] suggested by Shynk and Gooch [3]. As shown in the figure, the NLS IIR ADF converges faster than the RGN IIR ADF for which the optimum convergence factor α was chosen by trial and error to be 0.05.

It is known that, when the least mean squares (LMS) or self-orthogonalizing type algorithm is used for an IIR ADF, the optimum convergence factor is chosen heuristically in order for the filter to converge fast to the minimum steady state error. But, the learning speed of the proposed NLS IIR ADF is not sensitive to the value of the weighting factor λ in the practical range of minimum and maximum values (i.e., 0.9 – 1). In other words, λ may be chosen arbitrarily (e.g., 0.95) for the low-order case. It may be worthwhile to mention that, according to our experiment, while the NLS IIR ADF converged within 100 data samples for the system studied, the well-known LMS type IIR ADF required about 1500 samples to converge [2].

---

1 For a fair comparison we use a slightly modified RGN IIR ADF composed of only one subfilter with two poles.

We now consider the performance of the NLS IIR ADF for a high-order case which may be regarded as a generalized IIR ADF that models an arbitrary finite number of poles and zeros. One may note that, when a filter has more than two poles, the instability problem can become an important issue. To overcome such a problem, the high-order filter can be decomposed into parallel low-order subfilters using discrete Fourier transform (DFT) [3]. As a high-order case, we investigate a system whose transfer function is given by

$$G(z) = \frac{K(1+z^{-1})(1+1.41z^{-1}+z^{-2})}{(1-1.56z^{-1}+0.81z^{-2})(1-0.98z^{-1}+0.96z^{-2})} \tag{13}$$

where K is chosen such that the system output has unity power. Taking an 8-point DFT of input signal and decomposing the NLS IIR ADF into 8 frequency-bin subfilters (each of them has one pole and one zero), we obtained the *a priori* output estimate by summing the outputs of those subfilters. In Fig.3, the MSE is shown together with that of the frequency-domain RGN IIR ADF [3] under the same initial conditions (i.e., zero initial coefficient vectors). The stability of the two IIR ADF's was maintained by monitoring the poles of those subfilters. As seen in this figure, the learning speed of the NLS IIR ADF is slightly faster than that of the frequency-domain RGN IIR ADF, although the two ADF's have similar computational complexities. But, for a high-order system which has poles near the unit circle, such as the one being studied, the convergence speed of the NLS IIR ADF turned out to be sensitive to the weighting factor.

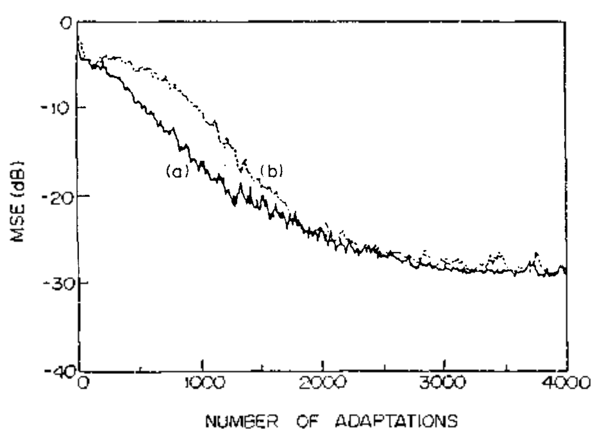

그림 3 극점이 두개이고 영점이 세개인 무한응답 적응 여파기의 수렴특성 (앙상블 평균을 위하여 독립적으로 100번 수행하였음)

Fig. 3 MSE of IIR ADF with 4 poles and 3 zeros (100 independent iterations were done for an ensemble average)

(a) NLS IIR ADF with $\lambda = 0.992$

(b) Frequency-domain RGN IIR ADF with $\alpha = 0.008$ (result of [3])

REFERENCES

[1] C.R. Johnson, Jr., "Adaptive IIR filtering : Current results and open issues," IEEE Trans. Inform. Theory, vol. IT-30, no.2, pp. 237-250, Mar. 1984.

[2] H. Fan and W.K. Jenkins, "A new adaptive IIR filter ," IEEE Trans. Circuits Syst., vol. CAS-33, no. 10, pp. 939-947, Oct. 1986.

[3] J.J. Shynk and R.P. Gooch, "Frequency-domain adaptive pole-zero filtering," Proc. IEEE, vol.73, no.10, pp. 1526-1528, Oct. 1985.

[4] G.C. Goodwin and K. S. Sin, Adaptive Filtering, Prediction and Control, Prentice-Hall, N.J., 1984.