'91 KACC 1991. 10. 22~24

A Numerically Efficient Adaptive Filter Algorithm with Varying Step Size by the Error

Byung-Eul Jun and Dong-Jo Park

Department of Electrical Engineering

Korea Advanced Institute of Science and Technology (KAIST)

Abstract – A numerically efficient modification of a variable step size LMS (Least Mean Squares) algorithm is proposed. This proposed algorithm is very simple in calculation and has a variable step size adjusted by the filter output error. Its additional computational burden with respect to the conventional LMS algorithm is only two multiplications, two substraction, an addition and some bit operations. In a simulation example, it is shown that the proposed algorithm has not only the faster convergence rate but also less misadjustments in the environment of highly nonstationary and correlated data .

I Introduction

The LMS(least mean square) Widrow-Hoff algorithm [1] is very simple in the number of calculations required for its update and robust in a number of practical applications. Therefore it has been useful in adaptation of the weight vector of adaptive FIR(finite impulse response) filters and in training of the multilayer perceptron, which is generally of high order. In these applications the LMS algorithm has worked well, but its convergence rate is fairly slow. The convergence rate of the LMS algorithm is dependent on the step size, i.e. the large step size results in the fast convergence. However the large step size also causes the algorithm to get large misadjustments of the filter weights, since the convergence rate and misadjustments have a trade-off relationship each other.

At this stage, the problem is that we want the fast convergence rate as well as little error residuals in the adaptive algorithm in spite of the trade-off relationship. Some modifications of the LMS Widrow-Hoff algorithm were proposed as solutions of the problem [2], [3], [4], [5]. A common method used in these algorithms is that the step size can be varied on going. In this paper another solution is proposed, which is very simple in calculation to update filter weights and has a variable step size adjusted by the filter output error. A similar idea on the step size adjustment was used in the Karni-Zeng algorithm [4]. But the calaulation of the proposed algorithm is more simple and its step size depends not on the norm of the gradient but on the filter output error.

II The Proposed Algorithm

The LMS Widrow-Hoff algorithm is

$$W(t+1) = W(t) - \mu\nabla(t),\ t = 0, 1, 2, \ldots, \quad (1)$$

where t is the time index in the discrete-time domain, $W \in R^M$ is the filter weight vector and μ is a constant step size. And the gradient $\nabla \in R^M$ is defined as

$$\nabla(t) = \frac{\partial \varepsilon^2(t)}{2\,\partial W} = -\varepsilon(t)X(t) \tag{2}$$

In equation (2), $X \in R^M$ is the filter input vector, and ε is the filter output error. The error is defined

$$\varepsilon(t) = d(t) - y(t) \tag{3}$$

$$y(t) = W^T(t)X(t), \tag{4}$$

where d is the desired output, y is the filter output and the superscript $(\cdot)^T$ denotes the transpose of the vector. In equation (1) the step size has a value within the interval

$$\mu_{min} < \mu < \mu_{max}. \tag{5}$$

In this relation (5), if the filter inputs are identically gaussian distributed with power σ^2, then the μ_{max} is given as [1]

$$\mu_{max} = \frac{2}{M\sigma^2}. \tag{6}$$

If the step size is kept large before the filter weight vector converges near to the optimal point and reduced appropriately as the vector converges, then the algorithm has the faster convergence property during the transient stage as well as less misadjustments of weights after convergence near to the optimal point. An implementation of this philosophy can be given as

$$\mu(t) = (\mu_{max} - \mu_{min})(1 - 2^{L(t)}) + \mu_{min}, \tag{7}$$

$$L(t) = -NINT(|\alpha\varepsilon(t)|), \tag{8}$$

where $NINT(\cdot)$ is defined as the nearest integer of $(\cdot)$, $|\cdot|$ is the absolute value and α is a weighting parameter of the filter output error. In equation (7), the L^{th} power of 2 can be calculated by bit shift operations in digital processors.

This simplicity in calculation makes the algorithm more practical from the implementation point of view. In that proposed algorithm, the calculation to update weight is very simple compared with the Karni-Zeng algorithm. Its additional computational burden with respect to the LMS Widrow-Hoff algorithm is only two multiplications, two substraction, an addition and some bit operations, but the Karni-Zeng algorithm additionally needs an exponential calculation and a calculation of the norm of the gradient, $\|\nabla(t)\|$. Another characteristic is that the step size is only dependent on the absolute value of the filter output error, which is different from the norm of the gradient in the Karni-Zeng algorithm.

III Simulations

To illustrate the convergence properties of the weight vector when employing the proposed algorithm, an adaptive linear combiner depicted on Fig. 1 is considered, where $r(k)$ is a white, zero-mean gaussian signal with variance σ^2 and $d(k)$ is a desired output signal defined as $d(k) = 2\cos(2\pi k/N)$.

Fig. 2 - Fig. 4 show simulation results of the proposed, the Karni-Zeng and the LMS Widrow-Hoff algorithm, which are plots of the ensemble average value at that time index of a twenty five executions. In these figures, convergence results of three algorithms are plotted on a figure easy to compare with. Environments for these simulations are as follows: the random noise variance σ^2 is 0.01, the weighting parameter α is 1 in the proposed algorithm and also in the Karni-Zeng algorithm. The maximum step size μ_{max} is 0.7 which is less than the one from equation (6), where the comments and results in [6], [7] were considered in order to choose the maximum step size μ_{max}. On the other hand the minimum step size μ_{min} is set to be 0.1.

And in the simulations we change the filter weights alternatively in order to test performance of the algorithms in the different phases of time varying environments. The N, samples per a signal cycle, is changed from 5 to 13 or vice versa periodically during these simulations. The change of N causes the MSE(mean square error) surface to be altered considerably. The eigenvalue spread, defined as λ_{max} / λ_{min} (where $\lambda_{max}[\lambda_{min}]$ is the maximum[minimum] eigenvalue of a matrix), of the input correlation matrix is about 2 when N is 5, and it is about 14 when N is 13. We call these situations phase 1 and phase 2 respectively as shown in the figures. This says that the simulated environments are highly nonstationary.

Fig. 2 shows variations of step sizes in the different phases where reduction of step sizes is apparent. The proposed and the Karni-Zeng algorithm worked better than the LMS Widrow-Hoff algorithm as shown in Fig. 3 and Fig. 4 because these algorithms can be adapted to the given nonstationary conditions via varying the step sizes by the filter output error. The proposed algorithm works as good as the Karni-Zeng algorithm in the phase 1, but in the phase 2 where the eigenvalue spread is large, the proposed algorithm performs better than the Karni- Zeng algorithm.

In the environment where the data are highly correlated, the eigenvalue spread is large. In this case the magnitude of the slope of the MSE surface along the eigenvector of the smallest eigenvalue is usually small in the transient stage while the error is big as shown in the literature [1]. Therefore the norm of the gradient is smaller than the absolute value of the error. The step size in the LMS Widrow-Hoff algorithm is set to be a fixed value 0.1 as the minimum step size in the proposed algorithm. This value of the step size is roughly around the adjusted one during the first stage of the proposed one.

IV Conclusion

The proposed algorithm is basically a modification of the Karni-Zeng algorithm. However its calculation to vary the step size is much simple and the adjustment method of it is more reasonable because, in the case of highly correlated data, the norm of gradient is less than the absolute value of the error. In order to adjust the step size, the proposed one does not need the exponential calculation and gradient norm calculation which are needed in the Karni-Zeng algorithm. Moreover the step size in the proposed algorithm depends on only the absolute error rounded to an integer so that the calculation becomes just bit shift operations. This simplicity makes the proposed algorithm easy to be implemented in discrete-time domain by digital processors.

References

[1] B. Widrow and D. Stearns, *Adaptive Signal Processing,* Englewood Cliffs, N.J. : Prentice-Hall, 1985.

[2] R. W. Harris, D. M. Chabries, and F. A. Bishop, "A Variable Step(VS) Adaptive Filter Algorithm," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-34, pp. 309-316, April 1986.

[3] R. A. Jacobs, "Increased Rates of Convergence Through Learning Rate Adaptation," *Neural Network,* vol.1, pp. 295-307, 1988.

[4] S. Karni and G. Zeng, "A New Convergence Factor for Adaptive Filters," *IEEE Trans. Circuits Syst.,* vol. CAS-36, pp. 1011-1012, July 1989.

[5] T. J. Shan and T. Kailath, "Adaptive Algorithms with an Automatic Gain Control Features," *IEEE trans. Circuits Syst.,* vol. CAS-35, pp. 122-127,

January 1988.

[6] A. Feuer and E. Weinstein, "Convergence Analysis of LMS Filters with Uncorrelated Gaussian Data," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 222-230, February 1985.

[7] V. A. Gholkar, "Mean Square Convergence Analysis of LMS Algorithm," Electron. Lett., vol. 26, pp. 1705-1706, 27th September 1990.

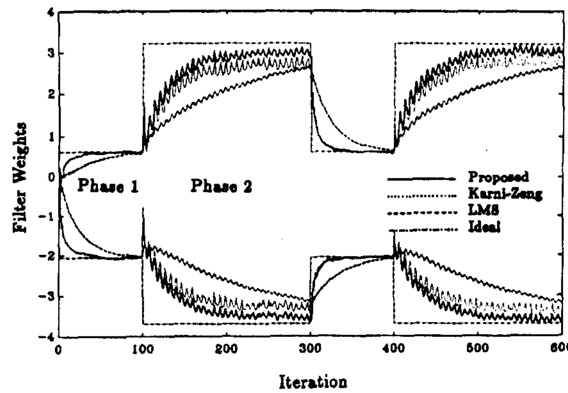

Fig. 3 Convergence Properties of Filter Weights

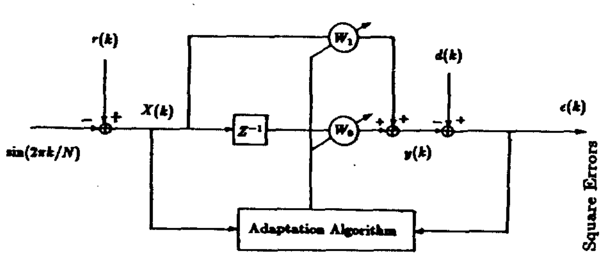

Fig. 1 An Adaptive Linear Combiner System

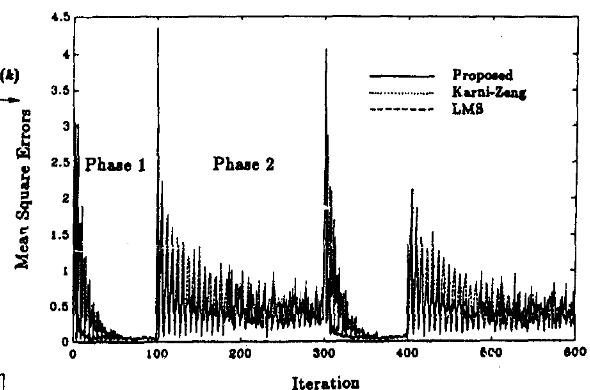

Fig. 4 Convergence Properties of Mean Square Errors

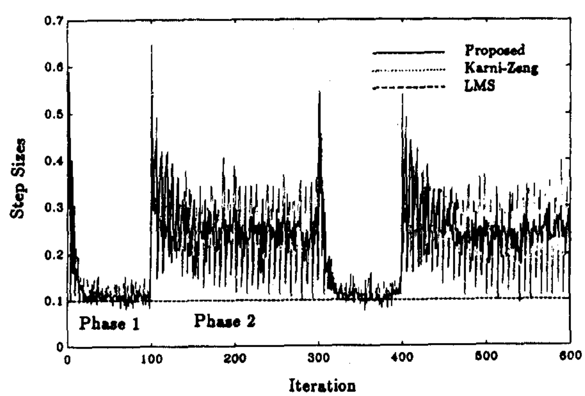

Fig. 2 Variations of Step Sizes