

An Approach to Variable Step-Size LMS Algorithm

Božo Krstajić, Ljubiša Stanković, Zdravko Uskoković

Abstract— A new approach to step size adaptation in the Variable Step-Size Least Mean Square (VS LMS) algorithm is proposed. This solution is based on the weighting coefficients bias/variance trade-off.

[3], [4]:

$$\overline{W}(k+1) = \overline{W}(k) + 2\mu e(k)\overline{X}(k) \quad (1)$$

I. INTRODUCTION

The LMS algorithm is the most popular one for implementation of adaptive filters [1], [2]. The VS LMS adaptive algorithms, [6] aim to improve the LMS algorithm performance. These algorithms use a different step size for each adaptive filter coefficient; the step size is adjusted individually as adaptation progresses. They are more efficient than the LMS algorithm for coefficients tracking in nonstationary environments [6]. The increase in complexity of implementation is relatively low.

Our approach for adaptation of step size is based on the investigation of the LMS weighting vector error in a nonstationary environment, due to both the effects of additive noise and weighting vector lag, [1], [2]. Minimization of the mean square deviation (MSD), i.e. the covariance of the weighting vector error, with respect to the step size, leads to the conclusion that the influences of additive noise and weighting vector lag are equal for the optimal step size [1], [2]. This condition is used in a specific statistical approach to produce the adaptive step size close to the optimal one, [3], [4], [5]. It is employed here to obtain the new formula for step size adaptation, based on the MSD minimization, which leads to a new type of the VS LMS algorithm. Its performance is illustrated by simulation results.

II. A NEW VS LMS ALGORITHM

In the LMS algorithm, the vector of the weighting coefficients is obtained from, [1], [2],

where μ is the step size, $\overline{X}(k)$ is the N -dimensional input signal vector, $e(k) = d(k) - y(k) = d(k) - \overline{W}^T(k)\overline{X}(k)$ is the output error, $d(k)$ is the reference signal and $y(k)$ is the filter output. The LMS adaptive filter tries to adjust a set of weighting coefficients so that $y(k)$ tracks $d(k) = \overline{W}_o^T(k)\overline{X}(k) + n(k)$, where $n(k)$ is a zero-mean Gaussian noise with the variance σ_n^2 and $\overline{W}_o(k) = \overline{W}_o(k-1) + \Delta\overline{W}_o(k-1)$ is time-varying optimal (Wiener) vector, where the covariance of $\Delta\overline{W}_o(k-1)$ is $\Theta(k) = \text{diag}(\theta, \theta, \dots, \theta)$, as in [2]. As shown in [2], $\overline{W}(k)$ can be decomposed as $\overline{W}(k) = E(\overline{W}(k)) + \overline{W}_n(k) + \overline{W}_l(k)$, where $E(\overline{W}(k))$ is the mean, $\overline{W}_n(k)$ and $\overline{W}_l(k)$ are the zero mean fluctuation of $\overline{W}(k)$ due to the additive noise and to the tracking lag, respectively. The MSD for the i -th weighting coefficient (MSD_i), obtained for step size μ_i , in steady state, is a sum of MSD's for each component, $MSD_i = MSD_{ib} + MSD_{in} + MSD_{il}$, where $MSD_{ib} = \theta/(4\mu_i\lambda_i)$ is the MSD due to lag bias, referred to as the squared weighting coefficient bias, i.e. bias_i^2 , $MSD_{in} = \mu_i\sigma_n^2$ is the MSD due to additive noise, referred to as the weighting coefficient variance σ_i^2 , and $MSD_{il} = \theta(N+1)/4$ is MSD due to lag variance, [2]. Since the MSD_{il} is independent on the step size μ_i , it will not affect our analysis, where we consider the abrupt changes of optimal vector. Thus, it is negligible with respect to the first two components of MSD_i , i.e. $MSD_i \cong \text{bias}_i^2 + \sigma_i^2$. As shown in [2], minimization of MSD_i with respect to μ_i gives the optimal step size $\mu_i^o = \sqrt{\theta/(4\lambda_i\sigma_n^2)}$, for which the MSD_{in} is equal to the MSD_{ib} , i.e. the

MSD optimality condition follows:

$$MSD_i n |_{\mu_i^o} = MSD_i b |_{\mu_i^o}, \quad \text{i.e. } bias_i^2 |_{\mu_i^o} = \sigma_i^2 |_{\mu_i^o} \quad (2)$$

Since the relation for μ_i^o is not applicable in practice, because one needs to know the input signal autocorrelation matrix eigenvalue (λ_i), the nature of optimal vector changes (θ) and noise variance (σ_n^2), we will here provide one possible way to estimate μ_i^o . Our approach does not require knowledge of λ_i and θ , while we estimate σ_n^2 by the weighting coefficient variance σ_i^2 .

Namely, note that the i -th weighting coefficient $W_i(k)$ assumes random values around the optimal $W_i^o(k)$ one with the $bias_i(k)$ and the variance σ_i^2 , related by, [3], [4], [5]:

$$|W_i(k) - W_i^o(k) - bias_i(k)| \leq \kappa \sigma_i, \quad (3)$$

where the above inequality holds with the probability $P(\kappa)$, dependent on κ . For example, for $\kappa = 2$ and a Gaussian distribution, this is satisfied with 95% probability. Let us now define the confidence intervals for random values $W_i(k)$:

$$C_i(k) = [W_i(k) - (\kappa + \Delta\kappa)\sigma_i, W_i(k) + (\kappa + \Delta\kappa)\sigma_i] \quad (4)$$

where the parameter $\Delta\kappa$ takes into account the $bias_i(k)$. Then, from (3) and (4) we can conclude that, as long as the $bias_i(k)$ is small, i.e. $|bias_i(k)| < \Delta\kappa\sigma_i$, the optimal value $W_i^o(k)$ belongs to the confidence interval $C_i(k)$, independently on the step size. It means that, for small bias, the confidence intervals (4) for different step sizes intersect. When, on the other hand, the bias becomes large, then the central positions of the intervals are far apart for different step sizes, and they do not intersect, [4], [5], [6].

Taking, as a criterion for the bias/variance trade-off, the condition that the bias and variance are of the same order of magnitude $|bias_i(k)| \approx \Delta\kappa\sigma_i$ we get the criterion for choosing the step size value. Namely, by using the MSD optimality condition (2) we take $\Delta\kappa \approx 1$.

Note that VS LMS coefficients updates are obtained from (1) if μ is replaced with $\bar{M}(k) = diag(\mu_0(k), \mu_1(k), \dots, \mu_{N-1}(k))$.

Aiming to optimize the step size $\mu_i(k)$ for the i -th weighting coefficient of VS LMS, we will compare its confidence interval with the one of the LMS with maximal allowed step size $\mu_{M \max}$, i.e. the one with best tracking of abrupt changes of optimal vector, [1], [2]. If we take that both algorithms start the k -th iteration with the same coefficients values from the $(k-1)$ -st iteration, then, taking into account (3), (4), and (1), the above comparison reduces to following inequality:

$$\begin{aligned} & 2 |e(k)x(k-i)| (\mu_{\max} - \mu_i(k)) \\ & \leq (\kappa + 1) \sigma_n^2 (\sqrt{\mu_{\max}} + \sqrt{\mu_i(k)}) \end{aligned} \quad (5)$$

Here we used the known fact that the variance for LMS is $\sigma_i^2 = \mu_i k \sigma_n^2$, [1], [2]. The best bias/variance ratio is obtained for the particular step size that turns (5) into an equality, thus producing the relation for the step size close to the optimal one:

$$\sqrt{\mu_i(k)} = \begin{cases} \sqrt{\mu_{\max}} - \frac{(\kappa+1)\sigma_n}{2a(k)} & \text{if } \mu_i(k) > \mu_{\min} \\ \sqrt{\mu_{\min}} & \text{if } \mu_i(k) \leq \mu_{\min} \end{cases} \quad (6)$$

where $a(k) = |e(k)x(k-i)|$, and μ_{\min} is the particular defined minimal step size value.

In our simulations we have estimated the noise power σ_n^2 by the weighting coefficient variance σ_i^2 , obtained by [5], [6]:

$$\sigma_i = median(|W_i(k) - W_i(k-1)|) / 0.675\sqrt{2}, \quad (7)$$

for $k = 1, 2, \dots, L$.

The above relation produces good estimates for all stationary cases, as well as for nonstationary ones, including abrupt changes of the optimal vector.

III. EXAMPLE AND CONCLUSION

The proposed New VS LMS (NVSS) algorithm is implemented in nonstationary environments, in a system identification setup. The algorithm performance is compared with the Harris VS LMS algorithm (HVSS) and the robust VS LMS (RVSS) algorithm [6]. Presented results are obtained by averaging over 200 independent runs, with number of weighting coefficients $N = 4$ and same order of unknown system. In the first 30 iterations the

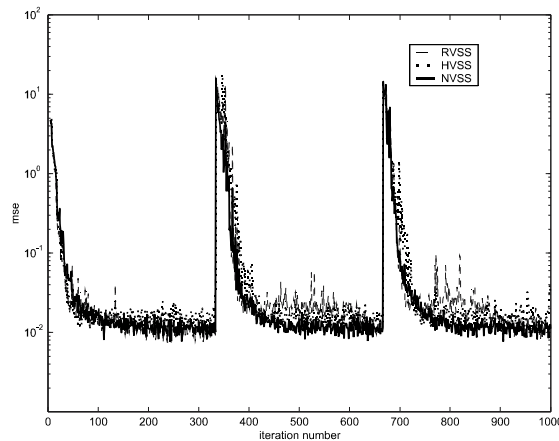


Fig. 1. Comparison of MSE for the considered algorithms

noise power was estimated according to (7) and our algorithm used μ_{max} . Other parameters values are: $\mu_{max}=0.1$, $\mu_{min}=0.001$, $SNR = 16dB$ (for all algorithms); $\kappa = 2$ (for the NVSS), $\alpha = 0.97$, $\beta = 0.99$, $\gamma = 0.5$ (for the RVSS) and $m_o=4$, $m_1=3$ (for the HVSS). Multiplying all the system coefficients by -1 at the 333 - th and the 666 - th iteration generates the abrupt changes of optimal vector. Figure 1 shows the mean square error (MSE) characteristics for each considered algorithm. In order to clearly compare the obtained results, we calculated the average $MSE(MSEa)$. It was $MSEa = 0.23164$, $MSEa = 0.29565$ and $MSEa = 0.29271$, for the NVSS, the RVSS and HVSS, respectively.

The proposed NVSS algorithm differs from other known VS LMS algorithms only in the criterion for the step size change. It enables one to get close to the MSD minimum, which cannot be reached due to the unknown signal nature, the introduced assumptions and to the inevitable estimation error for σ_n^2 . Presented simulation results and analysis, for the case of abrupt optimal vector changes, show advantages of the proposed solution with respect to other known algorithms. Note that various performed simulations, not included here, show that NVSS performs as well as the known VS LMS algorithms in other stationary and nonstationary cases.

IV. ACKNOWLEDGMENT

This work is supported by the Volkswagen Stiftung, Federal Republic of Germany.

REFERENCES

- [1] B. Widrow, J. M. McCool, M. G. Larmore and C. R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter", Proc. of IEEE, vol. 64, No. 8, pp. 1947-1951, August 1976.
- [2] K. C. Ho, "A study of two adaptive filters in tandem", IEEE Trans. SP., vol. 48, No. 6, pp. 1626-1636, June 2000.
- [3] B. Krstajic, L.J. Stankovic, Z. Uskokovic, "Combined adaptive system for identification of unknown systems with varying parameters in a noisy environment", 6th ICECS, Conf. Proc.; vol. II, 745-748, Pafos, Cyprus, 1999
- [4] B. Krstajic, Z. Uskokovic, L.J. Stankovic, "Identification of unknown systems with varying parameters in a noisy environment with a new VS LMS algorithm", MED'01, Conf. Proc., Dubrovnik, Croatia, June 2001.
- [5] L.J. Stankovic, V. Katkovnik, "Algorithm for the instantaneous frequency estimation using time-frequency distributions with variable window width", IEEE Sign. Proc. Letters, vol. 5, No. 9, pp 224-227, September 1998.
- [6] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: Analysis and simulations", IEEE Trans. SP., vol. 45, No. 3, pp.631-639, March 1997.