

# The Binormalized Data-Reusing LMS Algorithm

J. A. Apolinário Jr.<sup>†‡</sup>  
apolin@coe.ufrj.br

M. L. R. de Campos<sup>†</sup>  
campos@aquarius.ime.eb.br

P. S. R. Diniz<sup>‡</sup>  
diniz@coe.ufrj.br

<sup>†</sup>Instituto Militar de Engenharia  
Depto. Eng. Elétrica  
Pça Gal. Tibúrcio, 80  
Praia Vermelha, Rio de Janeiro, RJ  
22.290-270 — Brazil

<sup>‡</sup>Universidade Federal do Rio de Janeiro  
COPPE - Programa de Engenharia Elétrica  
P. O. Box 68504  
Rio de Janeiro, RJ  
21945-970 — Brazil

## Abstract

A new algorithm, the binormalized data-reusing least mean squares (LMS) algorithm is presented. The new algorithm has been found to converge faster than other LMS-like algorithms, such as the Normalized LMS algorithm and several data-reusing LMS algorithms, when the input signal is highly correlated. The computational complexity of this new algorithm is only slightly higher than a recently proposed normalized new data-reusing LMS algorithm.

## 1 Introduction

The least mean squares (LMS) algorithm is very popular and has been widely used due to its simplicity. Its convergence speed, however, is highly dependent on the eigenvalue spread of the input-signal autocorrelation matrix (conditioning number) [1, 2]. Alternative schemes which try to improve this performance at the cost of minimum additional computational complexity have been proposed and extensively discussed in the past [1, 3, 4].

The data-reusing LMS (DR-LMS) algorithm, which uses current desired and input signals repeatedly within each iteration is one among such schemes. It can be easily shown that in the limit of infinite data reuses per iteration the DR-LMS and the normalized LMS (NLMS) algorithms yield the same solution [5]. Performance can be further improved with the recently proposed normalized and unnormalized new data-reusing LMS (NNDR-LMS and UNDR-LMS) algorithms [5]. These algorithms reuse the data pair, namely desired and input signals, from previous iterations as well.

In reference [5], a graphical description of NNDR-LMS and UNDR-LMS algorithms was presented and it was shown that this new class of data-reusing algorithms had prospective better performance than the NLMS algorithm. The geometric description also showed why improvement is achieved when the number of reuses is increased. The new binormalized data-reusing LMS (BNDR-LMS) algorithm introduced here employs normalization on two orthogonal directions obtained from consecutive data pairs within each iteration. In all simulations carried out with colored input signals, the new algorithm presented faster convergence than all other

algorithms mentioned above (case of two data pairs).

This paper is organized as follows. Section 2 presents the LMS-like algorithms as well as a graphical illustration of their coefficient updating. Section 3 introduces the new BNDR-LMS algorithm. Section 4 contains the simulation results and Section 5 draws some conclusions.

## 2 LMS, DR-LMS, NLMS and NDR-LMS Algorithms

For the LMS algorithm, the coefficient vector  $\mathbf{w}$  is updated in the opposite direction of the gradient vector obtained from instantaneous squared output error, i.e.,

$$\mathbf{w}_{LMS}(k+1) = \mathbf{w}_{LMS}(k) - \mu \nabla_w [e^2(k)]$$

where

$$e(k) = d(k) - \mathbf{x}^T(k) \mathbf{w}_{LMS}(k)$$

is the output error,  $d(k)$  is the desired signal,  $\mathbf{x}(k)$  is the input-signal vector containing the  $N+1$  most recent input-signal samples, and  $\mu$  is the step size. The coefficient-updating equation is

$$\mathbf{w}_{LMS}(k+1) = \mathbf{w}_{LMS}(k) + \mu e(k) \mathbf{x}(k)$$

For the DR-LMS with  $L$  data reuses, the coefficients are updated as

$$\mathbf{w}_{i+1}(k) = \mathbf{w}_i(k) + \mu e_i(k) \mathbf{x}(k)$$

for  $i = 0, \dots, L$ ;

where

$$e_i(k) = d(k) - \mathbf{x}^T(k) \mathbf{w}_i(k),$$
$$\mathbf{w}_0(k) = \mathbf{w}_{DR-LMS}(k),$$

and

$$\mathbf{w}_{DR-LMS}(k+1) = \mathbf{w}_{L+1}(k).$$

Note that if  $L = 0$  these equations correspond to the LMS algorithm and that  $\mu$  is the step-size.

The NLMS algorithm normalizes the step-size such that the relation  $\mathbf{x}^T(k) \mathbf{w}_{NLMS}(k+1) = d(k)$  is always satisfied, i.e.,

$$\mathbf{w}_{NLMS}(k+1) = \mathbf{w}_{NLMS}(k) + \frac{e(k)}{\mathbf{x}^T(k) \mathbf{x}(k) + \epsilon} \mathbf{x}(k)$$

where  $\epsilon$  is a very small number used to avoid division by zero.

The NNDR-LMS algorithm is specified by the following relations

$$\mathbf{w}_{i+1}(k) = \mathbf{w}_i(k) + \frac{e_i(k)}{\mathbf{x}^T(k-i)\mathbf{x}(k-i) + \epsilon} \mathbf{x}(k-i)$$

for  $i = 0, \dots, L$ ;

where

$$e_i(k) = d(k) - \mathbf{x}^T(k)\mathbf{w}_i(k),$$

$$\mathbf{w}_0(k) = \mathbf{w}_{NNDR-LMS}(k),$$

and

$$\mathbf{w}_{NNDR-LMS}(k+1) = \mathbf{w}_{L+1}(k).$$

For the sake of comparison, our interest is in one single reuse such that  $L = 1$ . Fig. 1 illustrates geometrically the updating of the coefficient vector for a two-dimensional problem for all algorithms discussed above, starting with an arbitrary  $\mathbf{w}(k)$ .

Let  $\mathcal{S}(k)$  denote the hyperplane which contains all vectors  $\mathbf{w}$  such that  $\mathbf{x}^T(k)\mathbf{w} = d(k)$ . In a noise-free perfect-modeling situation,  $\mathcal{S}(k)$  contains the optimal coefficient vector,  $\mathbf{w}_o$ . Furthermore, it can be easily shown that  $\mathbf{x}(k)$  and, consequently,  $\nabla_{\mathbf{w}}[e^2(k)]$  are orthogonal to the hyperplane  $\mathcal{S}(k)$ .

The solution given by the DR-LMS algorithm,  $\mathbf{w}(k+1)$ , iteratively approaches  $\mathcal{S}(k)$  by following the direction given by  $\mathbf{x}(k)$  (see 3 in Fig. 1). This solution would reach  $\mathcal{S}(k)$  in the limit, as the number of data reuses goes to infinity [5]. The NLMS algorithm performs a line search to yield the solution  $\mathbf{w}(k+1) \in \mathcal{S}(k)$  in a single step (see 4 in Fig. 1). The algorithms presented in [5] use more than one hyperplane, i.e., data pair  $(\mathbf{x}, d)$ , in order to produce a solution  $\mathbf{w}(k+1)$  (see 5 and 6 in Fig. 1) that is closer to  $\mathbf{w}_o$  than the solution obtained with only the current data pair  $(\mathbf{x}(k), d(k))$ . For a noise-free perfect-modeling situation,  $\mathbf{w}_o$  is at the intersection of  $N+1$  hyperplanes constructed with linearly independent input-signal vectors. In this case, the orthogonal-projections algorithm [6] yields the solution  $\mathbf{w}_o$  in  $N+1$  iterations. This algorithm may be viewed as a normalized data-reusing orthogonal algorithm which utilizes  $N+1$  data pairs  $(\mathbf{x}, d)$ .

In the next section, the new binormalized data-reusing LMS algorithm will be described. This algorithm combines data reusing, orthogonal projections of two consecutive gradient directions, and normalization in order to achieve faster convergence when compared to other LMS-like algorithms. At each iteration, the BNDR-LMS yields the solution  $\mathbf{w}(k+1)$  which is at the intersection of hyperplanes  $\mathcal{S}(k)$  and  $\mathcal{S}(k-1)$  and at a minimum distance from  $\mathbf{w}(k)$  (see 7 in Fig. 1). The algorithm can also be viewed as a simplified version of the orthogonal projections algorithm which utilizes just two previous consecutive directions.

### 3 The BNDR-LMS Algorithm

In order to state the problem, we recall that the solution which belongs to  $\mathcal{S}(k)$  and  $\mathcal{S}(k-1)$  at a minimum distance

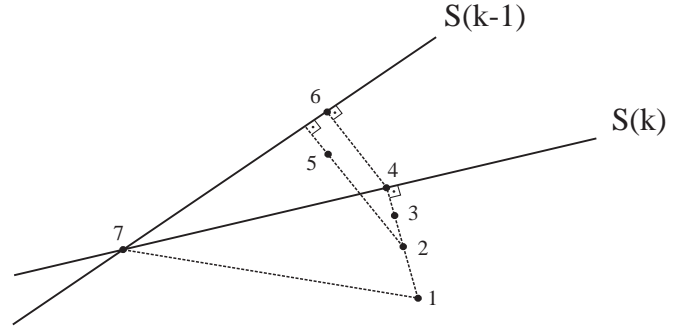


Figure 1: Updating the coefficient vector:

1.  $\mathbf{w}(k)$ ;
2.  $\mathbf{w}_{LMS}(k+1)$  (first step of DR-LMS and UNDR-LMS);
3.  $\mathbf{w}_{DR-LMS}(k+1)$ ;
4.  $\mathbf{w}_{NLMS}(k+1)$  (first step of NNDR-LMS);
5.  $\mathbf{w}_{UNDR-LMS}(k+1)$ ;
6.  $\mathbf{w}_{NNDR-LMS}(k+1)$ ;
7.  $\mathbf{w}_{BNDR-LMS}(k+1)$ .

from  $\mathbf{w}(k)$  is the one that solves

$$\min_{\mathbf{w}(k+1)} \|\mathbf{w}(k+1) - \mathbf{w}(k)\|^2$$

subjected to

$$\mathbf{x}^T(k)\mathbf{w}(k+1) = d(k)$$

and

$$\mathbf{x}^T(k-1)\mathbf{w}(k+1) = d(k-1)$$

The functional to be minimized is, therefore,

$$\begin{aligned} f[\mathbf{w}(k+1)] &= [\mathbf{w}(k+1) - \mathbf{w}(k)]^T [\mathbf{w}(k+1) - \mathbf{w}(k)] \\ &\quad + \lambda_1 [\mathbf{x}^T(k)\mathbf{w}(k+1) - d(k)] \\ &\quad + \lambda_2 [\mathbf{x}^T(k-1)\mathbf{w}(k+1) - d(k-1)] \end{aligned}$$

which, for linearly independent input-signal vectors  $\mathbf{x}(k)$  and  $\mathbf{x}(k-1)$ , has the unique solution

$$\mathbf{w}(k+1) = \mathbf{w}(k) + (-\lambda_1/2)\mathbf{x}(k) + (-\lambda_2/2)\mathbf{x}(k-1)$$

where

$$-\lambda_1/2 = \frac{num1}{den}$$

and

$$-\lambda_2/2 = \frac{num2}{den}$$

with:

$$\begin{aligned} num1 &= [d(k) - \mathbf{x}^T(k)\mathbf{w}(k)]\mathbf{x}^T(k-1)\mathbf{x}(k-1) \\ &\quad - [d(k-1) - \mathbf{x}^T(k-1)\mathbf{w}(k)]\mathbf{x}^T(k)\mathbf{x}(k-1) \\ num2 &= [d(k-1) - \mathbf{x}^T(k-1)\mathbf{w}(k)]\mathbf{x}^T(k)\mathbf{x}(k) \\ &\quad - [d(k) - \mathbf{x}^T(k)\mathbf{w}(k)]\mathbf{x}^T(k-1)\mathbf{x}(k) \\ den &= \mathbf{x}^T(k)\mathbf{x}(k)\mathbf{x}^T(k-1)\mathbf{x}(k-1) \\ &\quad - [\mathbf{x}^T(k)\mathbf{x}(k-1)]^2 + \epsilon \end{aligned}$$

where  $\epsilon$  is a small number used to avoid division by zero.

The BNDR-LMS algorithm is summarized in Table 1. This algorithm can be alternatively derived from a purely geometric reasoning. The first step is to reach a preliminary solution,  $\mathbf{w}_1(k)$ , which belongs to  $\mathcal{S}(k)$  and is at a minimum distance from  $\mathbf{w}(k)$ . This is achieved by the NLMS algorithm starting from  $\mathbf{w}(k)$ , i.e.,

$$\mathbf{w}_1(k) = \mathbf{w}(k) + \frac{e(k)}{\mathbf{x}^T(k)\mathbf{x}(k)}\mathbf{x}(k)$$

In the second step,  $\mathbf{w}_1(k)$  is updated in a direction orthogonal to the previous one, therefore belonging to  $\mathcal{S}(k)$ , until the intersection with  $\mathcal{S}(k-1)$  is reached. This is achieved by the NLMS algorithm starting from  $\mathbf{w}_1(k)$  and following the direction  $\mathbf{x}_1^\perp(k)$  which is the projection of  $\mathbf{x}(k-1)$  onto  $\mathcal{S}(k)$ .

$$\mathbf{w}(k+1) = \mathbf{w}_1(k) + \frac{e_1(k)}{\mathbf{x}_1^{\perp T}(k)\mathbf{x}_1^\perp(k)}\mathbf{x}_1^\perp(k)$$

where

$$\mathbf{x}_1^\perp(k) = \left[ \mathbf{I} - \frac{\mathbf{x}(k)\mathbf{x}^T(k)}{\mathbf{x}^T(k)\mathbf{x}(k)} \right] \mathbf{x}(k-1)$$

and

$$e_1(k) = d(k-1) - \mathbf{x}^T(k-1)\mathbf{w}_1(k)$$

The use of  $\mathbf{x}_1^\perp(k)$  obtained from  $\mathbf{x}(k-1)$  assures that the minimum-distance path is chosen.

Note that the requirement of linear independence of consecutive input-signal vectors  $\mathbf{x}(k)$  and  $\mathbf{x}(k-1)$ , necessary to ensure existence and uniqueness of the solution, is also manifested here. If  $\mathbf{x}(k)$  and  $\mathbf{x}(k-1)$  are linearly dependent, then we cannot find  $\mathbf{x}_1^\perp(k) \in \mathcal{S}(k)$ . This situation is avoided with  $\epsilon$  (we used  $10^{-5}$  in our experiments) in the algorithm.

Table 1: The Binormalized Data-Reusing LMS Algorithm.

BNDR-LMS
$\epsilon = \text{small value}$
for each $k$
{ $\mathbf{x}_1 = \mathbf{x}(k)$
$\mathbf{x}_2 = \mathbf{x}(k-1)$
$d_1 = d(k)$
$d_2 = d(k-1)$
$a = \mathbf{x}_1^T \mathbf{x}_2$
$b = \mathbf{x}_1^T \mathbf{x}_1$
$c = \mathbf{x}_2^T \mathbf{x}_2$
$d = \mathbf{x}_1^T \mathbf{w}(k)$
$e = \mathbf{x}_2^T \mathbf{w}(k)$
$den = bc - a^2 + \epsilon$
$A = (d_1c + ea - dc - d_2a)/den$
$B = (d_2b + da - eb - d_1a)/den$
$\mathbf{w}(k+1) = \mathbf{w}(k) + A\mathbf{x}_1 + B\mathbf{x}_2$
}

## 4 Simulation Results

In order to test the BNDR-LMS algorithm, simulations were carried out for a system identification problem. The

system order was  $N = 10$ , the input signal was correlated noise with a conditioning number around 55 and a input signal to observation noise ratio  $SNR = 150\text{dB}$ . The learning curves (MSE in dB) for the NLMS, the NNDR-LMS (one reuse) and the BNDR-LMS are depicted in Fig. 2, corresponding to an average of 200 realizations.

In this example we can clearly verify the superior performance of the BNDR-LMS algorithm in terms of speed of convergence when compared to the NLMS and the NNDR-LMS (with one single reuse) algorithms. Simulations for the conventional LMS algorithm and for the DR-LMS algorithm were also carried out for the same setup, but their performances were, as expected, inferior than that of the NLMS algorithm and the results were omitted from Fig. 2.

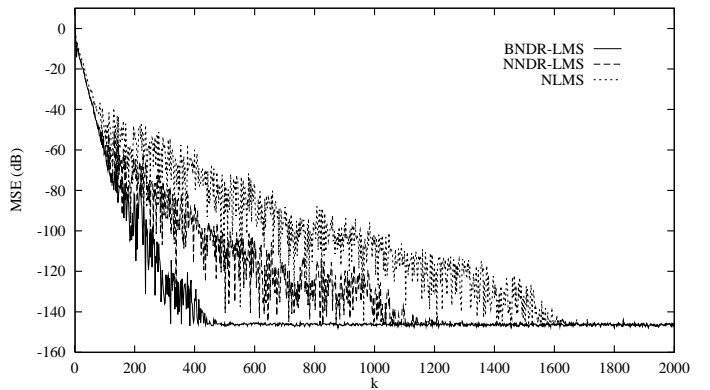


Figure 2: Learning curves of the following algorithms: NLMS, NNDR-LMS and BNDR-LMS.

In order to test the performance of the algorithms in terms of mean-square error after convergence, we measured the excess of MSE ( $\text{MSE} - \text{MSE}_{min}$ ) in dB. The  $\text{MSE}_{min}$  is the variance of the observation noise set equal to  $10^{-6}$  in this experiment. The results are summarized in Table 2 where we can also observe the excess of MSE in dB for a nonstationary environment. In this case, the observation noise was set to zero and the system (plant) coefficients varied according to  $\mathbf{w}(k) = \mathbf{w}(k-1) + \mathbf{v}$ , where  $\mathbf{v}$  is a random vector with zero mean and variance equal to  $10^{-6}$ . As we can see from Table 2, in both stationary and nonstationary environment, the BNDR-LMS algorithm performed closely to NLMS and NNDR-LMS algorithms.

Table 2: Excess Mean-Square Error.

Algorithm Type	$(\text{MSE} - \text{MSE}_{min})_{dB}$	
	Stationary	Nonstationary
NLMS	-59.09	-39.15
NNDR-LMS	-59.40	-39.42
BNDR-LMS	-58.60	-39.45

In terms of computational complexity, Table 3 shows the comparisons among these three algorithms. Note that  $p = N + 1$  is the number of coefficients.

Table 3: Comparison of computational complexity.

ALG.	ADD	MULT.	DIV.
NLMS	$3p-1$	$3p$	1
NNDR-LMS	$6p-2$	$6p$	2
BNDR-LMS	$7p+3$	$7p+2$	2

## 5 Conclusions

This paper introduced the BNDR-LMS algorithm which has a faster convergence than a number of other LMS-like algorithms when the input signal is highly correlated. A geometric interpretation of the algorithm was also provided showing that the coefficients are updated in two normalized steps following orthogonal directions. The relationship between the BNDR-LMS algorithm and the orthogonal-projections algorithm was clarified. Simulations carried out in a system identification application showed that the BNDR-LMS algorithm compared favorably with other LMS-like algorithms in terms of speed of convergence. Moreover, the more correlated is the input signal, the better the new algorithm performs. This improvement is clearly verified in cases of high signal to noise ratio.

## References

- [1] P.S.R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*. Kluwer Academic Press, Boston, MA, USA, 1997.
- [2] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, Englewood Cliffs, NJ, USA, 2nd edition, 1991.
- [3] S. Roy and J.J. Shynk, "Analysis of the data-reusing LMS algorithm," *32nd Midwest Symposium Circuits and Systems*, pp. 1127-1130, Urbana-Champaign, IL, USA, 1990.
- [4] D. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Transactions on Signal Processing*, vol. 41, pp. 2811-2825, Sept. 1993.
- [5] B.A. Schnauffer, *Practical Techniques for Rapid and Reliable Real-Time Adaptive Filtering*, Ph.D. Thesis, Urbana-Champaign, IL, USA, 1995.
- [6] G.C. Goodwin and K.S. Sin, *Adaptive Filtering Prediction and Control*. Prentice-Hall, Englewood Cliffs, NJ, USA, 1984.