Research article

Data selection with set-membership affine projection algorithm

Akram Zardadi*

Department of Mathematics, Payame Noor University (PNU), P.O. Box, 19395-4697, Tehran, Iran

* **Correspondence:** Email: azardadi1990@yahoo.com.

Abstract: In this paper, the set-membership affine projection (SM-AP) algorithm is utilized to censor non-informative data in big data applications. To this end, the probability distribution of the additive noise signal and the excess of mean-squared error (EMSE) in steady-state are employed in order to estimate the threshold parameter of the single threshold SM-AP (ST-SM-AP) algorithm aiming at attaining the desired update rate. Furthermore, by defining an acceptable range for the error signal, the double threshold SM-AP (DT-SM-AP) algorithm is proposed to detect very large errors due to the irrelevant data such as outliers. The DT-SM-AP algorithm can censor non-informative and irrelevant data in big data applications, and it can improve misalignment and convergence rate of the learning process with high computational efficiency. The simulation and numerical results corroborate the superiority of the proposed algorithms over traditional algorithms.

Keywords: adaptive filtering; set-membership filtering; affine projection; data censoring; big data; outliers

1. Introduction

Data redundancy is a ubiquitous feature in machine learning and big data applications, and it results in high computational load, energy consumption, processing time, and memory usage. By exploiting data redundancy, we can improve the learning performance and reduce the computational cost. An efficient approach to exploit data redundancy is through censoring non-informative data. Indeed, there are many works in literature benefiting from data censoring, such as censoring outliers in radar data [1], big data processing [2, 3], multi-sensor systems [4], sensor-centric data reduction [5], data censoring for energy-efficient communications [6], wireless sensor networks [7], data selective adaptive filters for sparse systems [8, 9], just to name a few. These works illustrate the advantages of censoring data in comparison with utilizing all data.

The set-membership filtering (SMF) is an efficient technique in dividing data into informative and non-informative data set [10]. In the SMF approach, instead of processing all data in the learning

process, we evaluate, select, then process data at each iteration. Indeed, the set-membership algorithms execute a new update whenever the output estimation error is larger than a predetermined value and incoming dataset contains enough innovation. Otherwise, the set-membership approach prevents the algorithm from implementing new updates; thus we will have a reduction in the computational burden. The most well known set-membership algorithms are the set-membership normalized least-mean-square [11, 12] and the set-membership affine projection [13] algorithms, where they have the benefits of the normalized least-mean-square and the affine projection algorithms, respectively, but they decrease the computational complexity through exploiting data redundancy and censoring data. Furthermore, there are many variants of the set-membership algorithms and their applications in the literature [14–22].

The SMF is an efficient approach for censoring data in big data applications, whereas due to the very large amount of data it would be more practical to determine beforehand the amount/percentage of data we desire to use in the learning process. In real applications, many restrictions can affect our ability to analyze the incoming data, such as limitations on available energy, time, memory, etc. Therefore, to surmount the restrictions and obtain the desired performance, we should be able to determine the informative incoming data. In this paper, we use the desired probability of updating of parameters in order to estimate the threshold parameter in the single threshold set-membership affine projection (ST-SM-AP) algorithm. Indeed, the function of the threshold parameter in the SM-AP algorithm is to censor non-informative data, thus estimating this parameter is crucial to censor data properly.

The set-membership algorithms, including the ST-SM-AP, censor incoming data based on the energy of the error signal. Indeed, if the energy of the error signal is larger than the threshold parameter, the ST-SM-AP algorithm will update the coefficients of the adaptive filter; otherwise they will remain unchanged. However, very high errors do not always show the presence of informative data. More precisely, sometimes we observe very high error signal because of the existence of some irrelevant data such as outliers and, in these cases, censoring data is more practical. Hence, we propose the double threshold set-membership affine projection (DT-SM-AP) algorithm which considers an acceptable range for the absolute value of the error signal to censor non-informative and irrelevant data and to prevent unnecessary updates. The DT-SM-AP algorithm implements the update whenever the absolute value of the error signal is between two threshold parameters [23]. Also, it worth to mention that some data-selective adaptive filtering algorithms have already been proposed in [21]; however, this work implemented an online threshold parameter computation; i.e., it updates the threshold parameter at any iteration. In our method, the threshold parameter is defined offline; i.e., it is computed before running the algorithm to determine the percentage of data we want to utilize to obtain the desired result.

This paper is organized as follows. Section 2 reviews the ST-SM-AP algorithm. Section 3 proposes the estimate of the threshold parameter in order to censor non-informative incoming data. The DT-SM-AP algorithm is introduced in Section 4. Simulations and conclusions are presented in Sections 5 and 6, respectively.

Notations: Scalars are denoted by lower case letters. Vectors (matrices) are represented by lowercase (uppercase) boldface letters. At iteration k, the optimum solution, the weight vector, and the input vector are denoted by $\mathbf{w}_o$, $\mathbf{w}(k)$, $\mathbf{x}(k) \in \mathbb{R}^{N+1}$, respectively, where N is the adaptive filter order. For a given iteration k, the error signal is defined as $e(k) \triangleq d(k) - \mathbf{w}^T(k)\mathbf{x}(k)$, where $d(k) \in \mathbb{R}$ is the desired signal and $(\cdot)^T$ stands for the vector and matrix transposition. Moreover, $P[\cdot]$ and $E[\cdot]$

denote the probability and the expected value operators, respectively.

2. The Single threshold set-membership affine projection algorithm

By taking advantage of old data, data-reusing algorithms can improve the convergence rate of the learning process, particularly when the input signal is correlated. The affine projection (AP) algorithm has been traditionally considered as the benchmark among data-reusing algorithms; however, the AP algorithm is not capable of exploiting data redundancy. By incorporating the set-membership technique to the AP algorithm, the single threshold set-membership affine projection (ST-SM-AP) algorithm has already been introduced [13] in order to reduce the computational burden of the AP algorithm. However, we intend to utilize the ST-SM-AP algorithm in order to censor data and exploit data redundancy in big data applications. First, let us define some necessary variables for the ST-SM-AP algorithm. Suppose that $\mathbf{x}(k)$ and $d(k)$ are the input vector and the desired signal, respectively, and the last $L + 1$ input vectors and desired signals are available. For a given iteration k, suppose that the input matrix $\mathbf{X}(k)$, the input vector $\mathbf{x}(k)$, the adaptive filter $\mathbf{w}(k)$, the desired vector $\mathbf{d}(k)$, the additive noise vector $\mathbf{n}(k)$, the constraint vector (CV) $\boldsymbol{\gamma}(k)$, and the error vector $\mathbf{e}(k)$ are described as follows

$$
\begin{aligned}
\mathbf{X}(k) &= [\quad \mathbf{x}(k) \quad \mathbf{x}(k-1) \quad \cdots \quad \mathbf{x}(k-L) \quad] \quad \in \quad \mathbb{R}^{(N+1)\times(L+1)}, \\
\mathbf{x}(k) &= [\quad x(k) \quad x(k-1) \quad \cdots \quad x(k-N) \quad]^T \quad \in \quad \mathbb{R}^{N+1}, \\
\mathbf{w}(k) &= [\quad w_0(k) \quad w_1(k) \quad \cdots \quad w_N(k) \quad]^T \quad \in \quad \mathbb{R}^{N+1}, \\
\mathbf{d}(k) &= [\quad d(k) \quad d(k-1) \quad \cdots \quad d(k-L) \quad]^T \quad \in \quad \mathbb{R}^{L+1}, \\
\mathbf{n}(k) &= [\quad n(k) \quad n(k-1) \quad \cdots \quad n(k-L) \quad]^T \quad \in \quad \mathbb{R}^{L+1}, \\
\boldsymbol{\gamma}(k) &= [\quad \gamma_0(k) \quad \gamma_1(k) \quad \cdots \quad \gamma_L(k) \quad]^T \quad \in \quad \mathbb{R}^{L+1}, \\
\mathbf{e}(k) &= [\quad e_0(k) \quad e_1(k) \quad \cdots \quad e_L(k) \quad]^T \quad \in \quad \mathbb{R}^{L+1},
\end{aligned}
\tag{2.1}
$$

where N and L are the adaptive filter order and the data-reuse factor, respectively. The entries of $\boldsymbol{\gamma}(k)$ should satisfy $|\gamma_i(k)| \leq \overline{\gamma}_1$, for $i = 0, 1, \cdots, L$, where $\overline{\gamma}_1 \in \mathbb{R}_+$ is the upper bound for the magnitude of the error signal. Moreover, The error vector $\mathbf{e}(k)$ is defined as $\mathbf{e}(k) = \mathbf{d}(k) - \mathbf{X}^T(k)\mathbf{w}(k)$.

We can now represent the recursion rule of the ST-SM-AP algorithm by [13]

$$
\mathbf{w}(k+1) = \begin{cases} \mathbf{w}(k) + \mathbf{X}(k)\left[\mathbf{X}^T(k)\mathbf{X}(k) + \delta\mathbf{I}\right]^{-1}(\mathbf{e}(k) - \boldsymbol{\gamma}(k)) & \text{if } |e(k)| > \overline{\gamma}_1, \\ \mathbf{w}(k) & \text{otherwise,} \end{cases}
\tag{2.2}
$$

where $\delta \in \mathbb{R}_+$ and $\mathbf{I} \in \mathbb{R}^{(L+1)\times(L+1)}$ are a regularization factor and the identity matrix, respectively, and $\delta\mathbf{I}$ is added to $\mathbf{X}^T(k)\mathbf{X}(k)$ in order to avoid numerical problems in the matrix inversion. In the following section, we intend to introduce an approach to estimate $\overline{\gamma}_1$ such that it leads to the desired update rate in big data applications.

3. Estimating $\overline{\gamma}_1$ in the ST-SM-AP algorithm

In this section, we obtain an estimate for $\overline{\gamma}_1$ in the ST-SM-AP algorithm for online censoring in streaming big data applications. In the presence of data redundancy and streaming data, it is desirable to obtain an acceptable solution by using a predetermined percentage of data instead of processing

all received data. The percentage of data we want to use gives us the update rate of the ST-SM-AP algorithm. Therefore, for a given predetermined update rate, we intend to estimate the threshold $\overline{\gamma}_1$ so that the update rate of the ST-SM-AP algorithm does not exceed the predetermined update rate. In other words, for a given $0 < p < 1$ and considering Equation (2.2), we intend to estimate $\overline{\gamma}_1$ such that

$$P[|e(k)| > \overline{\gamma}_1] = p. \tag{3.1}$$

In the above equation, the predetermined update rate p represents the percentage of data we would like to utilize in the learning process. Note that the responsibility of $\overline{\gamma}_1$ is to choose the most informative data, by considering p, for the learning process.

If we know the probability distribution of the error signal $e(k)$, then we can calculate the suitable value of $\overline{\gamma}_1$. Generally, we do not have access to the probability distribution of the error signal. However, using the central limit theorem [24], when the adaptive filter order is sufficiently large, the error signal $e(k)$ will have a zero-mean Gaussian distribution [22]. Therefore, by utilizing the probability distribution of $n(k)$ in (3.1), we will be able to calculate the proper value of $\overline{\gamma}_1$.

Note that the error signal $e(k)$ is obtained by subtracting the output of the adaptive filter from the desired signal, thus we have

$$e(k) = d(k) - \mathbf{x}^T(k)\mathbf{w}(k) = \mathbf{x}^T(k)\mathbf{w}_o + n(k) - \mathbf{x}^T(k)\mathbf{w}(k) = \mathbf{x}^T(k)[\mathbf{w}_o - \mathbf{w}(k)]^T + n(k) = \widetilde{e}(k) + n(k), \tag{3.2}$$

where $\widetilde{e}(k)$ is the noiseless error signal. Furthermore, we know that the ST-SM-AP algorithm is robust [25], thus $\|E[\mathbf{w}_o - \mathbf{w}(k)]\|^2 < \infty$, for all $k \in \mathbb{N}$. It means that the ST-SM-AP algorithm never diverges. In general, we have $E[\mathbf{w}_o - \mathbf{w}(k)] \approx \mathbf{0}$ in the steady-state, where $\mathbf{0}$ stands for the zero vector.

A notable example for the probability distribution of the additive noise signal is the zero-mean Gaussian noise with variance σ_n^2. Therefore, assuming this important case, we are capable of estimating the threshold $\overline{\gamma}_1$. As we know, the noiseless error signal $\widetilde{e}(k)$ is uncorrelated with the additive noise signal $n(k)$, thus using (3.2) we get

$$E[e(k)] = E[\widetilde{e}(k)] + E[n(k)] = 0, \tag{3.3}$$

$$\text{Var}[e(k)] = E[\widetilde{e}^2(k)] + \sigma_n^2. \tag{3.4}$$

The value of $E[\widetilde{e}^2(k)]$ is called the excess of mean-squared error (EMSE), and for the ST-SM-AP algorithm in the steady-state is described by [27]

$$E[\widetilde{e}^2(k)] = \frac{(L+1)[\sigma_n^2 + \overline{\gamma}_1^2 - 2\overline{\gamma}_1\sigma_n^2\rho]p}{[(2-p) - 2(1-p)\overline{\gamma}_1\rho]}\left(\frac{1-a}{1-a^{L+1}}\right), \tag{3.5}$$

where

$$\rho = \sqrt{\frac{2}{\pi(2\sigma_n^2 + \frac{1}{L+1}\overline{\gamma}_1^2)}}, \tag{3.6}$$

$$a = [1 - p + 2p\overline{\gamma}_1\rho](1 - p). \tag{3.7}$$

As can be seen in Equation (3.5), in order to compute $E[\widetilde{e}^2(k)]$ we need $\overline{\gamma}_1$, whereas our target is to estimate $\overline{\gamma}_1$. Therefore, initially, in the steady-state, we consider that the value of $E[\widetilde{e}^2(k)]$ is equal

to zero and the distribution of $e(k)$ is identical to the distribution of $n(k)$. It means that, for the first moment, $E[e(k)] = 0$, $Var[e(k)] = \sigma_n^2$, and the distribution of $e(k)$ is the zero-mean Gaussian with variance σ_n^2. Thus for a given p, using the relation (3.1), we can compute the initial estimate of $\overline{\gamma}_1$. In other words, we have

$$P[|e(k)| > \overline{\gamma}_1] = P[e(k) < -\overline{\gamma}_1] + P[e(k) > \overline{\gamma}_1] = p, \tag{3.8}$$

since the Gaussian distribution is symmetric, we get

$$P[e(k) > \overline{\gamma}_1] = \frac{p}{2}. \tag{3.9}$$

Hence, we should compute $\overline{\gamma}_1$ by equation

$$\int_{\overline{\gamma}_1}^{\infty} \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp(-\frac{r^2}{2\sigma_n^2}) dr = \frac{p}{2}, \tag{3.10}$$

where $\exp(\cdot)$ stands for the exponential function. Thus, for a given update rate p, we can find the initial estimate of $\overline{\gamma}_1$ using the standard normal distribution table.

Then, by having the initial estimate of $\overline{\gamma}_1$ at hand, we should replace it in Equations (3.5), (3.6), and (3.7) to calculate $E[\widetilde{e}^2(k)]$. Ultimately, by having $E[\widetilde{e}^2(k)]$, we will compute the variance of $e(k)$ utilizing Equation (3.4). Thus, the distribution of the error signal is the zero-mean Gaussian with variance $\sigma_e^2 = Var[e(k)] = E[\widetilde{e}^2(k)] + \sigma_n^2$ and, by utilizing this distribution probability, we can attain a better estimate for $\overline{\gamma}_1$. Indeed, the estimate of $\overline{\gamma}_1$ can be obtained from equation

$$\int_{\overline{\gamma}_1}^{\infty} \frac{1}{\sqrt{2\pi\sigma_e^2}} \exp(-\frac{r^2}{2\sigma_e^2}) dr = \frac{p}{2}, \tag{3.11}$$

where we can achieve the new estimate of $\overline{\gamma}_1$ by standard normal distribution table. we should remind that the selected update rate p describes a loose relative importance of the innovation caused by the new incoming input and desired data set.

4. The double threshold set-membership affine projection algorithm

In the previous section, a process to estimate the threshold parameter of the ST-SM-AP algorithm has been addressed such that the update rate of the algorithm does not exceed a predetermined value p. The ST-SM-AP algorithm avoids updating the adaptive filter parameters when the absolute value of the error signal is less than $\overline{\gamma}_1$. Indeed, the ST-SM-AP algorithm assumes all incoming data set with absolute value error larger than $\overline{\gamma}_1$ as innovative information; however, this is not always true, particularly in the presence of outliers. In this section, we intend to introduce the double threshold set-membership affine projection (DT-SM-AP) algorithm to censor non-informative data and exploit data redundancy, also to avoid outliers effect, acquisition system saturation, and impulsive noise.

For a given iteration k, the ST-SM-AP algorithm update $\mathbf{w}(k)$ if $|e(k)| > \overline{\gamma}_1$, i.e., it updates $\mathbf{w}(k)$ when $|e(k)|$ is sufficiently large. However, there are iterations in which the error signal is very large, and it does not contain new information. In fact, very large error signal might be happened due to some irrelevant information at incoming data such as outliers, system saturation, and impulsive noise.

Hence, when the absolute value of the error signal is very large, it would be more practical to avoid new update in adaptive filter coefficients. To this end, we recommend to define an acceptable range for the absolute value of the error signal and to adopt a lower threshold $\overline{\gamma}_1$ and an upper threshold $\overline{\gamma}_2$. Then, at each iteration k, if the incoming data set results in an error $e(k)$ such that $\overline{\gamma}_1 < |e(k)| < \overline{\gamma}_2$, we implement a new updates since it means that the incoming data set brought about some innovation; otherwise we avoid new update since the incoming data brought about non-innovative or irrelevant information.

Finally, the update equation of the DT-SM-AP algorithm can be given by

$$\mathbf{w}(k+1) = \begin{cases} \mathbf{w}(k) + \mathbf{X}(k)\left[\mathbf{X}^T(k)\mathbf{X}(k) + \delta\mathbf{I}\right]^{-1}(\mathbf{e}(k) - \boldsymbol{\gamma}(k)) & \text{if } \overline{\gamma}_1 < |e(k)| < \overline{\gamma}_2, \\ \mathbf{w}(k) & \text{otherwise.} \end{cases} \quad (4.1)$$

The task of $\overline{\gamma}_1$ is to ignore the incoming data set without enough innovation, thus we can utilize the strategy described in the previous section in order to estimate $\overline{\gamma}_1$. However, $\overline{\gamma}_2$ is responsible for detecting irrelevant incoming data set, such as outliers, thus depending on the applications we should adopt a large enough value for $\overline{\gamma}_2$.

5. Simulations

In this section, we have applied the AP, the ST-SM-AP and DT-SM-AP algorithms to a system identification problem. For the ST-SM-AP algorithm, the threshold parameter $\overline{\gamma}_1$ is estimated employing the strategy explained in Section 3 in order to update the ST-SM-AP algorithm with the desired update rate p. The unknown system $\mathbf{w}_o$ has 10 coefficients, and they are drawn from the normal distribution. We have tested the results for three different input signals, namely binary phase-shift keying (BPSK), zero-mean white Gaussian noise with unit variance (WGN), and the first-order autoregressive signal (AR(1)) produced by $x(k) = 0.95x(k-1) + m(k)$, where $m(k)$ is a WGN signal. The signal-to-noise ratio is 20 dB; i.e., the additive noise signal has variance $\sigma_n^2 = 0.01$. The regularization parameter δ is set to be 10^{-12}, and the initialization for the adaptive filter is adopted as $\mathbf{w}(0) = [0 \cdots 0]^T$. The constraint vector (CV) is selected as the simple choice CV [25, 26]. The simple choice CV is defined as $\gamma_0(k) = \overline{\gamma}_1 \frac{e(k)}{|e(k)|}$ and $\gamma_i(k) = e_i(k)$ for $i = 1, \cdots, L$. Moreover, the step-size parameter for the AP algorithm is represented at the legend of each figure. The number of iterations is 5×10^4, and the learning curves and the update rates are the averages of outcomes of 100 independent ensembles.

5.1. Scenario 1

In this scenario, we have utilized the estimated threshold parameter $\overline{\gamma}_1$ in the ST-SM-AP algorithm in order to achieve the desired results by the given update rates $p = 0.15, 0.30$, and 0.45. The estimated threshold parameters for $L = 0, 1, 2, 3, 4$ and $p = 0.15, 0.30, 0.45$ are obtained by using the argument described in Section 3, and they are listed in Table 1. Furthermore, the update rates of the ST-SM-AP algorithm using these threshold parameters and different input signals in 5×10^4 iterations are presented in Table 2. As can be seen, by utilizing the estimated $\overline{\gamma}_1$s, for all input signals the resulting update rates presented in Table 2 are close to the given values of p, such that in all cases the difference between the amount of resulting update rate and the given value of p is less than 0.02 (2 percent). Therefore,

by estimating $\overline{\gamma}_1$ for a given update rate p, we have suitably censored non-informative incoming data such that, with high precision, we have controlled the amount of data we intend to use in the learning process.

Table 1. The value of estimated $\overline{\gamma}_1$ for different values of p and L

p \ L	0	1	2	3	4
0.15	0.1661	0.1692	0.1716	0.1736	0.1753
0.30	0.1199	0.1222	0.1244	0.1266	0.1292
0.45	0.0871	0.0900	0.0962	0.0998	0.1037

Table 2. The resulting update rates using the estimated $\overline{\gamma}_1$ for different values of p and L for the ST-SM-AP algorithm with different input signals

Input signal	p \ L	0	1	2	3	4
BPSK	0.15	0.1339	0.1331	0.1337	0.1334	0.1332
	0.30	0.2920	0.2909	0.2933	0.2935	0.2949
	0.45	0.4407	0.4419	0.4441	0.4460	0.4474
WGN	0.15	0.1341	0.1329	0.1338	0.1332	0.1354
	0.30	0.2982	0.2960	0.2967	0.2982	0.3003
	0.45	0.4439	0.4452	0.4456	0.4474	0.4483
AR(1)	0.15	0.1451	0.1445	0.1463	0.1475	0.1484
	0.30	0.3007	0.3016	0.3025	0.3034	0.3046
	0.45	0.4513	0.4526	0.4522	0.4541	0.4557

Moreover, Figures 1(a), 1(b), and 1(c) depict the MSE learning curves of the AP and the ST-SM-AP algorithms when the input signals are BPSK, WGN, and AR(1), respectively. Also, the data reuse factor is chosen to be $L = 2$. In these figures, the desired update rate is $p = 0.15$, thus the estimated $\overline{\gamma}_1$ is computed as 0.1716. For the AP algorithm, two different step-sizes 0.1 and 0.9 are adopted. In Figures 1(a), 1(b), and 1(c), the update rates of the ST-SM-AP algorithm are 0.1337, 0.1338, and 0.1463, respectively, as presented in Table 2. On the one hand, we can observe that when the step-size of the AP algorithm is large, the AP algorithm can reach the steady-state as fast as the ST-SM-AP algorithm; however, the MSE of the ST-SM-AP algorithm is lower than that of the AP algorithm. On the other hand, we can see that for small step-size the AP algorithm can attain the same MSE of the ST-SM-AP algorithm; however, the convergence rate of the AP algorithm degrades significantly. Therefore, the ST-SM-AP algorithm can reach low MSE with a high convergence rate as compared to the AP algorithm. Furthermore, it is worth mentioning that the ST-SM-AP algorithm has extremely lower computational complexity in comparison with the AP algorithm since it censors non-informative incoming data and avoids unnecessary updates, whereas the AP algorithm updates the adaptive filter coefficients at every iteration. Hence, the ST-SM-AP algorithm has the great potential to reduce the computational burden in big data applications.
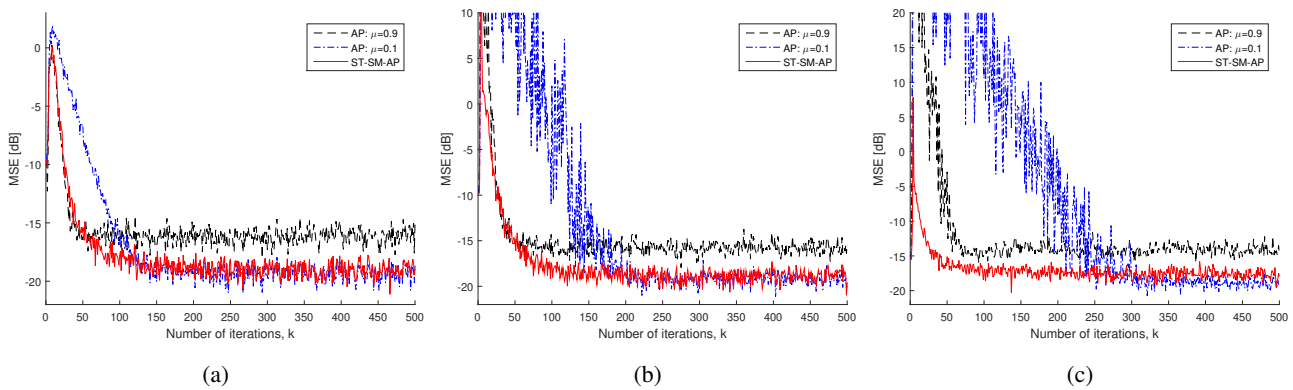
Figure 1. The MSE learning curves of the AP and the ST-SM-AP algorithms for $L = 2$ and $\overline{\gamma}_1 = 0.1716$ considering: (a) BPSK input signal; (b) WGN input signal; (c) AR(1) input signal.

5.2. Scenario 2

In this scenario, we have implemented the AP, the ST-SM-AP, and the DT-SM-AP algorithms in order to identify the unknown system $\mathbf{w}_o$ when the desired signal contains an outliers signal. The parameters of the mentioned algorithms are set identical to those of Scenario 1. The outliers signal is added to the reference signal, and it is produced by a Bernoulli process, which takes 1 with probability 0.05, multiplying uniformly distributed random numbers belong to the interval $(0, 50)$. For the DT-SM-AP algorithm, we adopt $\overline{\gamma}_2 = 1$ in order to detect outliers and avoid updating adaptive filter coefficients for irrelevant incoming data. By utilizing the threshold parameters $\overline{\gamma}_1$s presented in Table 1 and $\overline{\gamma}_2 = 1$, the update rates of the DT-SM-AP algorithm for different data reuse factors L and different input signals are given in Table 3. Similarly to Scenario 1, we can observe that in the presence of outliers signal, by utilizing the estimated $\overline{\gamma}_1$s, for all input signals the resulting update rates are given in Table 3 are close to the given value of p. Indeed, the DT-SM-AP algorithm detected outliers signal and non-informative data effectively and avoided unnecessary updates when receiving insignificant and irrelevant incoming data. It is worthwhile to notice that, in Table 3, the difference between the values of resulting update rates and their corresponding values of p is less than 0.02 (2 percent); therefore, it shows that the estimated $\overline{\gamma}_1$s and $\overline{\gamma}_2$ have censored the incoming data as we wanted.

Furthermore, Figures 2(a), 2(b), and 2(c) illustrate the misalignment curves of the AP, the ST-SM-AP, and the DT-SM-AP algorithms for the BPSK, WGN, and AR(1) input signals, respectively. For all mentioned algorithms, the data reuse factor L is selected to be 2. Also, the desired update rate is $p = 0.15$, therefore, the estimated $\overline{\gamma}_1$ is adopted equal to 0.1716 as given in Table 1, and $\overline{\gamma}_2 = 1$ in order to detect outliers and avoid unnecessary updates. Similarly to Scenario 1, we have adopted two step-sizes 0.1 and 0.9 in the AP algorithm, so that the small step-size results in low MSE and low convergence rate but the large step-size leads to high MSE and high convergence speed. In Figures 2(a), 2(b), and 2(c), the update rates of the DT-SM-AP algorithm are 0.1348, 0.1349, and 0.1458, respectively, as described in Table 3; however, the update rates of the ST-SM-AP algorithm are high values 0.8692, 0.8734, and 0.8985, respectively, because of unnecessary updates in the presence of outliers. As can be seen in these figures, the AP algorithm using neither large nor small step-sizes can reach the misalignment of the DT-SM-AP algorithm. Moreover, the misalignment and convergence

speed of the ST-SM-AP algorithm are the same as those of the AP algorithm using the large step-size 0.9. Therefore, in the presence of outliers, neither ST-SM-AP nor AP algorithms can attain the superb performance of the DT-SM-AP algorithm. Indeed, in the presence of outliers, the DT-SM-AP algorithm outperforms the ST-SM-AP and the AP algorithms by obtaining lower misalignment, lower update rate, and higher convergence rate.

Table 3. The resulting update rates using the estimated $\overline{\gamma}_1$s and $\overline{\gamma}_2 = 1$ for different values of p and L for the DT-SM-AP algorithm with different input signals

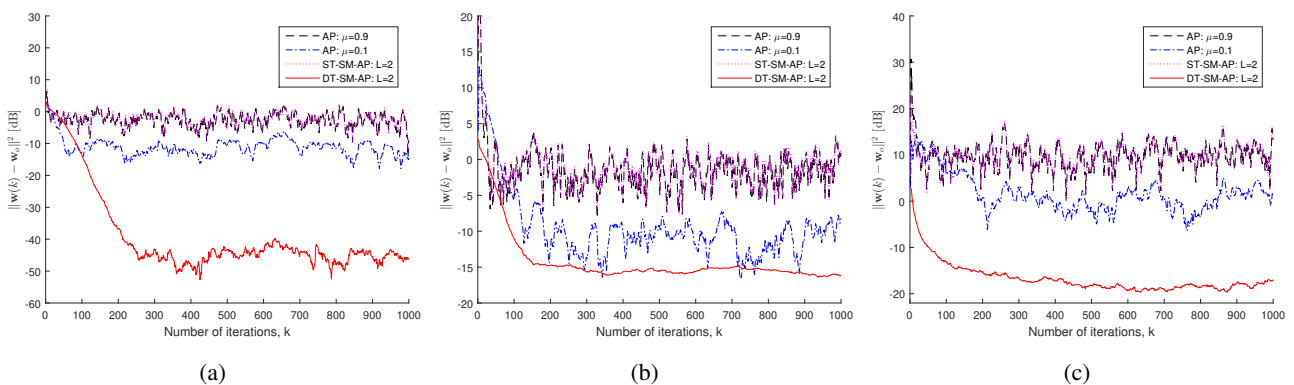| Input signal	L p	0	1	2	3	4
BPSK	0.15	0.1347	0.1342	0.1348	0.1341	0.1339
	0.30	0.2943	0.2934	0.2941	0.2944	0.2951
	0.45	0.4426	0.4435	0.4448	0.4473	0.4472
WGN	0.15	0.1352	0.1351	0.1349	0.1357	0.1363
	0.30	0.2973	0.2966	0.2974	0.2989	0.2996
	0.45	0.4448	0.4454	0.4469	0.4475	0.4487
AR(1)	0.15	0.1459	0.1452	0.1458	0.1471	0.1486
	0.30	0.29992	0.3009	0.3015	0.3026	0.3034
	0.45	0.4507	0.4518	0.4514	0.4529	0.4540

Figure 2. The misalignment curves of the AP, the ST-SM-AP, and the DT-SM-AP algorithms for $L = 2$, $\overline{\gamma}_1 = 0.1716$, and $\overline{\gamma}_2 = 1$ considering: (a) BPSK input signal; (b) WGN input signal; (c) AR(1) input signal.

6. Conclusions

In this paper, the single threshold set-membership affine projection (ST-SM-AP) and the double threshold set-membership affine projection (DT-SM-AP) algorithm have been proposed in order to exploit data redundancy and censor non-informative and irrelevant data aiming at improving the learning process. For this purpose, we have estimated the threshold parameter to attain the desired update rate in the proposed algorithms. The estimate of the threshold parameter has been obtained

with the help of the distribution probability function of the additive noise signal and the excess mean-squared error in steady-state of the algorithm. The ST-SM-AP algorithm avoids updating the adaptive filter coefficients when the absolute value of the output estimation error is smaller than the estimated threshold. However, the two thresholds in the DT-SM-AP algorithm prevent the algorithm from updating adaptive filter parameters when the absolute value of the error signal is outside the range defined by the thresholds. The numerical results indicate that the ST-SM-AP and the DT-SM-AP algorithms outperform the conventional affine projection algorithm.

Conflict of interest

The author declares no conflicts of interest in this paper.

References

1. Han S, De Maio S, Carotenuto V, et al. (2018) Censoring outliers in radar data: an approximate ML approach and its analysis. *IEEE T Aero Elec Sys* 55: 534–546.

2. Diniz PSR, Yazdanpanah H (2017) Data censoring with set-membership algorithms. In: *IEEE Global Conference on Signal and Information Processing (GlobalSIP 2017)*, Montreal, Canada, 121–125.

3. Zhu H, Qian H, Luo X, et al. (2018) Adaptive queuing censoring for big data processing. *IEEE Signal Proc Let* 25: 610–614.

4. Zheng Y, Niu R, Varshney PK (2014) Sequential bayesian estimation with censored data for multi-sensor systems. *IEEE T Signal Proces* 62: 2626–2641.

5. Msechu EJ, Giannakis GB (2012) Sensor-centric data reduction for estimation with WSNs via censoring and quantization. *IEEE T Signal Proces* 60: 400–414.

6. Fernández-Bes J, Arroyo-Valles R, Cid-Sueiro J (2011) Cooperative data censoring for energy-efficient communications in sensor networks. In: *IEEE International Workshop on Machine Learning for Signal Processing*, Santander, Spain, 1–6.

7. Msechu EJ, Giannakis GB (2011) Decentralized data selection for MAP estimation: a censoring and quantization approach. In: *14th International Conference on Information Fusion*, Chicago, IL, USA, 1–8.

8. Yazdanpanah H, Diniz PSR, Lima MVS (2016) A simple set-membership affine projection algorithm for sparse system modeling. In: *24th European Signal Processing Conference (EUSIPCO 2016)*, Budapest, Hungary, 1798–1802.

9. Yazdanpanah H, Diniz PSR (2017) Recursive least-squares algorithms for sparse system modeling. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2017)*, New Orleans, LA, USA, 3879–3883.

10. Diniz PSR (2013) *Adaptive Filtering: Algorithms and Practical Implementation*, 4th edition, New York, USA, Springer.

11. Gollamudi S, Nagaraj S, Kapoor S, et al. (1998) Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size. *IEEE Signal Proc Let* 5: 111–114.

12. Gollamudi S, Kapoor S, Nagaraj S, et al. (1998) Set-membership adaptive equalization and updator-shared implementation for multiple channel communications systems. *IEEE T Signal Proces* 46: 2372–2385.

13. Werner S, Diniz PSR (2001) Set-membership affine projection algorithm. *IEEE Signal Proc Let* 8: 231–235.

14. Yazdanpanah H, Diniz PSR (2017) New trinion and quaternion set-membership affine projection algorithms. *IEEE T Circuits-II* 64: 216–220.

15. Diniz PSR, Yazdanpanah H (2016) Improved set-membership partial-update affine projection algorithm. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016)*, Shanghai, China, 4174–4178.

16. Takahashi N, Yamada I (2009) Steady-state mean-square performance analysis of a relaxed set-membership NLMS algorithm by the energy conservation argument. *IEEE T Signal Proces* 57: 3361–3372.

17. Bhotto MZA, Antoniou A (2012) A robust constrained set-membership affine-projection adaptive-filtering algorithm. *IEEE T Signal Proces* 60: 73–81.

18. Deller JR (1989) Set-membership identification in digital signal processing. *IEEE ASSP Magazine* 6: 4–20.

19. Nagaraj S, Gollamudi S, Kapoor S, et al. (1999) BEACON: an adaptive set-membership filtering technique with sparse updates, *IEEE T Signal Proces* 47: 2928–2941.

20. Yazdanpanah H, Lima MVS, Diniz PSR (2016) On the robustness of the set-membership NLMS algorithm. In: *9th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2016)*, Rio de Janeiro, Brazil, 1–5.

21. Diniz PSR (2018) On Data-Selective Adaptive Filtering. *IEEE T Signal Proces* 66: 4239–4252.

22. Lima MVS, Diniz PSR (2013) Steady-state MSE performance of the set-membership affine projection algorithm. *Circ Syst Signal Pr* 32: 1811–1837.

23. Diniz PSR, Braga RP, Werner S (2006) Set-membership affine projection algorithm for echo cancellation. In: *International Symposium on Circuits and Systems (ISCAS 2006)*, Island of Kos, Greece.

24. Papoulis A (1991) *Probability, Random Variables, and Stochastic Processes*, 3rd edition, McGraw Hill, New York, USA.

25. Yazdanpanah H, Lima MVS, Diniz PSR (2017) On the robustness of set-membership adaptive filtering algorithms. *EURASIP J Adv Sig Pr* 2017: 72.

26. Martins WA, Lima MVS, Diniz PSR, et al. (2017) Optimal constraint vectors for set-membership affine projection algorithms. *Signal Process* 134: 285–294.

27. Lima MVS, Diniz PSR (2010) Steady-state analysis of the set-membership affine projection algorithm. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2010)*, Dallas, USA, 3802–3805.