Gholamreza Karamali
Akram Zardadi
Hamid Reza Moradi

DATA CENSORING WITH SET-MEMBERSHIP AFFINE PROJECTION ALGORITHM

Abstract *In this work, we use the single-threshold and double-threshold set-membership affine projection algorithm to censor non-informative and irrelevant data in big data problems. For this purpose, we employ the probability distribution function of the additive noise in the desired signal and the excess of the mean-squared error (EMSE) in steady-state to evaluate the threshold parameter of the single -threshold set-membership affine projection (ST-SM-AP) algorithm intending to obtain the desired update percentage. In addition, we propose the double-threshold set-membership affine projection (DT-SM-AP) algorithm to detect very large errors caused by unrelated data (such as outliers). The DT--SM-AP algorithm is capable of censoring non-informative and unrelated data in big data problems, and it will promote the misalignment and convergence speed of the learning procedure with low computational complexity. The synthetic examples and real-life experiments substantiate the superior performance of the proposed algorithms as compared to traditional algorithms.*

Keywords adaptive filtering, machine learning, data censoring, big data

Citation Computer Science 21(1) 2020: 43–57

1. Introduction

Big data is a growing area that means large volumes of structured, semi-structured, and unstructured data that poses a challenging job to be processed by using conventional approaches and databases. It is a technique for acquainted decision-making utilizing analytical methods to explain any data set that is massive enough that it needs the use of high-level programming skills and techniques to turn the data into an asset for a company [9]. Data abundance is an omnipresent characteristic in machine learning, adaptive filtering, and big data applications, and it leads to high computational burden, wastes of energy and time, and memory usage. By employing data redundancy, we can develop the learning achievement and decrease the computational load. A practical method for using data redundancy is by examining non-informative data. There are various papers benefiting from data censoring, such as censoring outliers in radar data [8], big data processing [5, 30], multi-sensor systems [29], sensor-centric data reduction [15], wireless sensor networks [14], and data selective adaptive filters for sparse systems [7, 20, 23]. These works explain the benefits of censoring data as compared to using all of the data.

Set-membership filtering (SMF) is a useful approach in apportioning data into informative and non-informative data sets [6]. In this technique, we evaluate, choose, and process the data at each iteration instead of assessing all of the data in the learning process. SMF algorithms implement a new update whenever the error is larger than a decided value and an incoming dataset comprises sufficient innovation. Contrarily, the SMF method stops the algorithm from realizing new updates; consequently, we will experience a decrease in computational resources. The most famous SMF algorithms are the set-membership normalized least-mean-square [1, 28] and the set-membership affine projection [19] algorithms, where they reduce computational costs by exploiting data redundancy. In addition, there are many variants of SMF algorithms in the literature [2, 4, 7, 11, 13, 17, 18, 22, 24].

SMF is a practical technique for censoring data in big data problems; however, owing to the massive volume of data, it would be more efficient to previously manage the volume/portion of data we want to use in the learning process. In real problems, various limitations can affect our capability to process incoming data, like restrictions on the available energy, time, memory, etc. Hence, to defeat the limitations and obtain the preferred outcome, we must be able to define the informative incoming data. In this work, by reviewing [27], we utilized the sought-after probability of updating coefficients to approximate the threshold parameter in the single-threshold set-membership affine projection (ST-SM-AP) algorithm to censor non-informative data.

SMF algorithms (including ST-SM-AP) control incoming data based on the absolute value of the error signal. In other words, if the energy of the error is greater than the threshold, the ST-SM-AP algorithm updates the adaptive coefficients of the system; otherwise, they remain unmodified. Very large errors do not always demonstrate the existence of informative data. To be more accurate, sometimes we perceive very large error due to the presence of some unrelated data (such as outliers); in these situations, censoring the data is more efficient. Thus, we introduce the double-threshold set-membership affine projection (DT-SM-AP) algorithm that assumes a satisfactory range for the absolute value of the error to cut non-informative and irrelevant data and avoid undesirable updates. The DT-SM-AP algorithm executes the update when the absolute value of the error is between two threshold parameters [3].

This work is organized as follows. Section 2 reviews the set-membership filtering approach. Section 3 describes the ST-SM-AP algorithm. Section 4 introduces the approximate of the threshold parameter to censor non-informative incoming data. Section 5 proposes the DT-SM-AP algorithm. Simulations and real-life experiments are presented in Sections 6. Finally, conclusions are drawn in Section 7.

Notations: Scalars are presented as lower-case letters, and vectors (matrices) are denoted by lower-case (upper-case) boldface letters. At given iteration k, the optimum solution, weight vector, and input vector are represented by $\mathbf{w}_o$, $\mathbf{w}(k)$, and $\mathbf{x}(k) \in \mathbb{R}^{N+1}$, respectively, where N is the adaptive filter order. For a given iteration k, the error is described by $e(k) \triangleq d(k) - \mathbf{w}^T(k)\mathbf{x}(k)$, where $d(k) \in \mathbb{R}$ is the desired signal, and $(\cdot)^T$ shows for the vector and matrix transposition. Furthermore, $P[\cdot]$ and $E[\cdot]$ stand for the probability and expected value operators, respectively. Moreover, $\mathbf{0}$ stands for the zero vector.

2. Set-membership filtering

The SMF approach proposed in [6,26] is suitable for adaptive filtering problems that are linear in parameters. Thus, for a given input signal vector $\mathbf{x}(k) \in \mathbb{R}^{N+1}$ at iteration k and filter coefficients $\mathbf{w} \in \mathbb{R}^{N+1}$, the output signal of the filter is obtained by

$$y(k) = \mathbf{w}^T\mathbf{x}(k) \tag{1}$$

where $\mathbf{x}(k) = [x_0(k)\ x_1(k)\ \cdots\ x_N(k)]^T$ and $\mathbf{w} = [w_0\ w_1\ \cdots\ w_N]^T$.

For a desired signal sequence $d(k)$, estimation error sequence $e(k)$ is computed as

$$e(k) = d(k) - y(k) \tag{2}$$

The SMF criterion aims at estimating parameter $\mathbf{w}$ such that the magnitude of the estimation output error is upper-bounded by a constant $\overline{\gamma} \in \mathbb{R}_+$ for all possible pairs $(\mathbf{x}, d)$. If the value of $\overline{\gamma}$ is suitably selected, there are various valid estimates for $\mathbf{w}$. The threshold is usually chosen based on *a priori* information about the sources of uncertainty. Note that any $\mathbf{w}$ leading to an output estimation error with a magnitude smaller than $\overline{\gamma}$ is an acceptable solution. Hence, we obtain a set of filters rather than a single estimate.

Let us denote as $\mathcal{S}$ the set comprised of all possible pairs $(\mathbf{x}, d)$. We want to find $\mathbf{w}$ such that $|e| = |d - \mathbf{w}^T\mathbf{x}| \leq \overline{\gamma}$ for all $(\mathbf{x}, d) \in \mathcal{S}$. Therefore, *feasibility set* Θ will be defined as

$$\Theta \triangleq \bigcap_{(\mathbf{x},d)\in\mathcal{S}} \{\mathbf{w} \in \mathbb{R}^{N+1} : |d - \mathbf{w}^T\mathbf{x}| \leq \overline{\gamma}\} \tag{3}$$

so that the SMF criterion can be stated as finding $\mathbf{w} \in \Theta$.

In the case of online applications, we do not have access to all members of $\mathcal{S}$. Thus, we consider the practical case in which only measured data is available and develop iterative techniques. Suppose that set of data pairs $\{(\mathbf{x}(0), d(0)), (\mathbf{x}(1), d(1)), \cdots, (\mathbf{x}(k), d(k))\}$ is available, and define *constraint set* $\mathcal{H}(k)$ at time instant k as

$$\mathcal{H}(k) \triangleq \{\mathbf{w} \in \mathbb{R}^{N+1} : |d(k) - \mathbf{w}^T\mathbf{x}(k)| \leq \overline{\gamma}\} \tag{4}$$

Also, define *exact membership set* $\psi(k)$ as the intersection of the constraint sets from the beginning (i.e., the first iteration) to iteration k, or

$$\psi(k) \triangleq \bigcap_{i=0}^{k} \mathcal{H}(i) \tag{5}$$

Then, Θ can be iteratively estimated via the exact membership set since $\lim_{k\to\infty} \psi(k) = \Theta$.

Figure 1 shows the geometrical interpretation of the SMF principle. The boundaries of the constraint sets are hyperplanes, and $\mathcal{H}(k)$ corresponds to the region between the parallel hyperplanes in the parameter space. The exact membership set represents a polytope in the parameter space. The volume of $\psi(k)$ decreases for each k in which pairs $(\mathbf{x}(k), d(k))$ bring about some innovation. Note that $\Theta \subset \psi(k)$ for all k, since Θ is the intersection of all possible constraint sets.

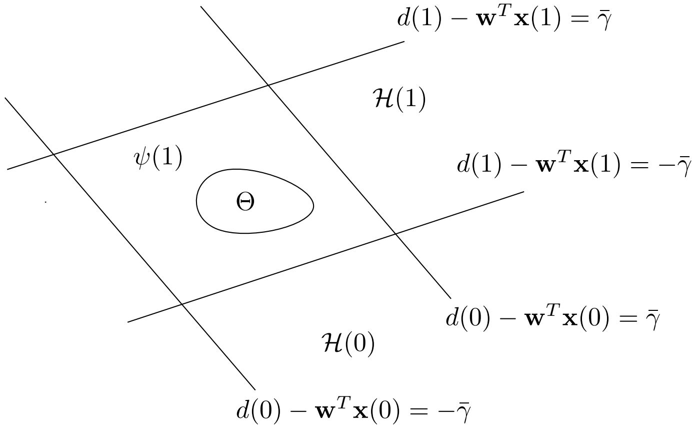
$d(1) - \mathbf{w}^T\mathbf{x}(1) = \bar{\gamma}$
$\mathcal{H}(1)$
$\psi(1)$
$d(1) - \mathbf{w}^T\mathbf{x}(1) = -\bar{\gamma}$
Θ
$d(0) - \mathbf{w}^T\mathbf{x}(0) = \bar{\gamma}$
$\mathcal{H}(0)$
$d(0) - \mathbf{w}^T\mathbf{x}(0) = -\bar{\gamma}$

Figure 1. SMF geometrical interpretation in parameter space $\psi(1)$

The target of set-membership adaptive filtering is to adaptively obtain an estimate that belongs to the feasibility set. The simplest method is to calculate a point estimate using, for example, the information provided by $\mathcal{H}(k)$ similar to the set-membership normalized least-mean-square algorithm or several previous $\mathcal{H}(k)$s (like in the SM-AP algorithm).

3. ST-SM-AP algorithm

Data-reusing algorithms can promote the convergence velocity of the learning procedure, especially if the input signal is correlated. For the last decades, the affine projection (AP) algorithm has been conventionally assumed to be the benchmark among data-reusing algorithms, whereas the AP algorithm is not able to take advantage of data redundancy. By joining the SM approach with the AP algorithm, the single-threshold set-membership affine projection (ST-SM-AP) algorithm has been proposed [19] to decrease the computational load of the AP algorithm. However, we want to employ the ST-SM-AP algorithm to censor data and use data abundance in big data problems. Let us introduce some basic variables for the ST-SM-AP algorithm. Assume that $\mathbf{x}(k)$ and $d(k)$ are the input vector and the desired signal, respectively, and we have access to the last $L+1$ $\mathbf{x}(k)$ and $d(k)$. At a given iteration k, assume that input matrix $\mathbf{X}(k)$, input vector $\mathbf{x}(k)$, adaptive filter $\mathbf{w}(k)$, desired vector $\mathbf{d}(k)$, additive noise vector $\mathbf{n}(k)$, constraint vector (CV) $\boldsymbol{\gamma}(k)$, and error vector $\mathbf{e}(k)$ are defined by

$$\begin{array}{rcllllllc}
\mathbf{X}(k) & = & [& \mathbf{x}(k) & \mathbf{x}(k-1) & \cdots & \mathbf{x}(k-L)] &] & \in & \mathbb{R}^{(N+1)\times(L+1)} \\
\mathbf{x}(k) & = & [& x(k) & x(k-1) & \cdots & x(k-N) &]^T & \in & \mathbb{R}^{N+1} \\
\mathbf{w}(k) & = & [& w_0(k) & w_1(k) & \cdots & w_N(k) &]^T & \in & \mathbb{R}^{N+1} \\
\mathbf{d}(k) & = & [& d(k) & d(k-1) & \cdots & d(k-L) &]^T & \in & \mathbb{R}^{L+1} \\
\mathbf{n}(k) & = & [& n(k) & n(k-1) & \cdots & n(k-L) &]^T & \in & \mathbb{R}^{L+1} \\
\boldsymbol{\gamma}(k) & = & [& \gamma_0(k) & \gamma_1(k) & \cdots & \gamma_L(k) &]^T & \in & \mathbb{R}^{L+1} \\
\mathbf{e}(k) & = & [& e_0(k) & e_1(k) & \cdots & e_L(k) &]^T & \in & \mathbb{R}^{L+1}
\end{array} \tag{6}$$

in which N and L are the adaptive filter order and data-reuse parameter, respectively. The components of $\boldsymbol{\gamma}(k)$ must satisfy $|\gamma_i(k)| \leq \overline{\gamma}_1$, for $i = 0, 1, \cdots, L$, in which $\overline{\gamma}_1 \in \mathbb{R}_+$ is the upper limit for the absolute value of the error. In addition, error vector $\mathbf{e}(k)$ is described by $\mathbf{e}(k) = \mathbf{d}(k) - \mathbf{X}^T(k)\mathbf{w}(k)$.

We can now characterize the update equation of the ST-SM-AP algorithm by [19]:

$$\mathbf{w}(k+1) = \begin{cases} \mathbf{w}(k) + \mathbf{X}(k)\Big[\mathbf{X}^T(k)\mathbf{X}(k) + \delta\mathbf{I}\Big]^{-1}(\mathbf{e}(k) - \boldsymbol{\gamma}(k)) & \text{if } |e(k)| > \overline{\gamma}_1, \\ \mathbf{w}(k) & \text{otherwise} \end{cases} \tag{7}$$

where $\delta \in \mathbb{R}_+$ and $\mathbf{I} \in \mathbb{R}^{(L+1)\times(L+1)}$ are a regularization parameter and the identity matrix, respectively, and $\delta\mathbf{I}$ is summed with $\mathbf{X}^T(k)\mathbf{X}(k)$ to prevent singular matrix inversion. In the next section, we will propose a strategy to approximate $\overline{\gamma}_1$ so that it results in the solicited update rate in big data problems.

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

By reviewing [27], we approximate $\overline{\gamma}_1$ in the ST-SM-AP algorithm for online censoring in flowing big data problems in this section. In streaming data (and when we have data abundance), it is helpful to attain a satisfactory solution by adopting a predetermined portion of data rather than processing all of the acquired data. Thus, given a predetermined update rate, we want to approximate threshold parameter $\overline{\gamma}_1$ such that the update rate of the ST-SM-AP algorithm does not pass the decided update percentage. This means that, for a given $0 < p < 1$ and assuming the recursion rule (7), we want to measure $\overline{\gamma}_1$ such that

$$\mathrm{P}[|e(k)| > \overline{\gamma}_1] = p \tag{8}$$

Note that, by considering p, $\overline{\gamma}_1$ is responsible for selecting the most informative data in the learning procedure.

To compute the suitable value of $\overline{\gamma}_1$, we must have access to the probability distribution of error signal $e(k)$. In general, the probability distribution of the error signal is not available. However, when the adaptive filter order is large enough, error signal $e(k)$ has a zero-mean Gaussian distribution [11]. Hence, by employing the probability distribution of $n(k)$ in (8), the suitable value of $\overline{\gamma}_1$ can be computed. Also, in [25], it has been shown that the ST-SM-AP algorithm is robust; therefore, $\|\mathrm{E}[\mathbf{w}_o - \mathbf{w}(k)]\|^2 < \infty$ for all $k \in \mathbb{N}$ and (generally) $\mathrm{E}[\mathbf{w}_o - \mathbf{w}(k)] \approx \mathbf{0}$ in the steady-state.

In many real-life problems, the probability distribution of the noise is the zero-mean Gaussian noise with variance σ_n^2. Hence, by considering this distribution, we can evaluate threshold $\overline{\gamma}_1$. Defining the noiseless error signal by $\widetilde{e}(k) = \mathbf{x}^T(k)[\mathbf{w}_o - \mathbf{w}(k)]^T$, we know that $\widetilde{e}(k)$ is uncorrelated with $n(k)$; thus, we get [27].

$$\mathrm{E}[e(k)] = \mathrm{E}[\widetilde{e}(k)] + \mathrm{E}[n(k)] = 0 \tag{9}$$

$$\mathrm{Var}[e(k)] = \mathrm{E}[\widetilde{e}^2(k)] + \sigma_n^2 \tag{10}$$

The excess of mean-squared error (EMSE) for the ST-SM-AP algorithm in the steady-state is given by [10].

$$\mathrm{E}[\widetilde{e}^2(k)] = \frac{(L+1)[\sigma_n^2 + \overline{\gamma}_1^2 - 2\overline{\gamma}_1\sigma_n^2\rho]p}{[(2-p) - 2(1-p)\overline{\gamma}_1\rho]}\Big(\frac{1-a}{1-a^{L+1}}\Big) \tag{11}$$

where

$$\rho = \sqrt{\frac{2}{\pi(2\sigma_n^2 + \frac{1}{L+1}\overline{\gamma}_1^2)}} \tag{12}$$

$$a = [1 - p + 2p\overline{\gamma}_1\rho](1-p) \tag{13}$$

Using Equation (11) to obtain $\mathrm{E}[\widetilde{e}^2(k)]$, we require $\overline{\gamma}_1$; thus, we first assume that $\mathrm{E}[e(k)] = 0$, $\mathrm{Var}[e(k)] = \sigma_n^2$, and the distribution of $e(k)$ is the zero-mean Gaussian with variance σ_n^2. Therefore, for a given p, the initial approximation of $\overline{\gamma}_1$ can be obtained by [27].

$$\int_{\overline{\gamma}_1}^{\infty} \frac{1}{\sqrt{2\pi\sigma_n^2}}\exp(-\frac{r^2}{2\sigma_n^2})dr = \frac{p}{2} \tag{14}$$

Afterwards, by using the acquired approximation of $\overline{\gamma}_1$, we substitute it into Equations (11), (12) and (13) to compute $\mathrm{E}[\widetilde{e}^2(k)]$. Thus, we can compute the variance of $e(k)$ by Equation (10); the distribution of the error would be a zero-mean Gaussian with variance $\sigma_e^2 = \mathrm{Var}[e(k)] = \mathrm{E}[\widetilde{e}^2(k)] + \sigma_n^2$. Therefore, the improved estimate for $\overline{\gamma}_1$ can be attained by [27].

$$\int_{\overline{\gamma}_1}^{\infty} \frac{1}{\sqrt{2\pi\sigma_e^2}}\exp(-\frac{r^2}{2\sigma_e^2})dr = \frac{p}{2} \tag{15}$$

5. DT-SM-AP algorithm

The ST-SM-AP algorithm updates the adaptive coefficients when the absolute value of the error is greater than $\overline{\gamma}_1$. In fact, the ST-SM-AP algorithm considers all incoming data with absolute value errors that are larger than $\overline{\gamma}_1$ as innovation, but this can be wrong (especially when there are outliers). In other words, a very large error can occur because of some irrelevant information from the incoming data, like outliers, system saturation, impulsive noise, etc. Thus, we review the double-threshold set--membership affine projection (DT-SM-AP) algorithm here to cut data without new information and to evade irrelevant data.

The idea of the DT-SM-AP algorithm is to avoid a new update when the absolute value of the error is very large. Thus, we assume an acceptable range for the error signal by choosing lower and upper threshold $\overline{\gamma}_1$ and $\overline{\gamma}_2$, respectively. Then, if $\overline{\gamma}_1 < |e(k)| < \overline{\gamma}_2$, we execute a new update; otherwise, we avoid a new update. Therefore, the recursion rule of the DT-SM-AP algorithm can be described by [27].

$$\mathbf{w}(k+1) = \begin{cases} \mathbf{w}(k) + \mathbf{X}(k)\Big[\mathbf{X}^T(k)\mathbf{X}(k) + \delta\mathbf{I}\Big]^{-1}(\mathbf{e}(k) - \boldsymbol{\gamma}(k)) & \text{if } \overline{\gamma}_1 < |e(k)| < \overline{\gamma}_2, \\ \mathbf{w}(k) & \text{otherwise} \end{cases} \tag{16}$$

Note that $\overline{\gamma}_1$ can be estimated using the strategy presented in the previous section. Moreover, the function of $\overline{\gamma}_2$ is to detect irrelevant incoming data; thus, depending on the applications, we can adopt a sufficient large value for $\overline{\gamma}_2$.

6. Experimental results

In this section, we use the AP, ST-SM-AP, and DT-SM-AP algorithms in numerical examples and real-life problems under system identification scenarios. We compute threshold parameter $\overline{\gamma}_1$ by using the approach proposed in Section 4 when the desired update rate is given as p. The unknown system is denoted by $\mathbf{w}_o$, and it is of order 9; i.e., it has ten coefficients. The coefficients of $\mathbf{w}_o$ are drawn from the Gaussian distribution with zero mean and unit variance. Three different input signals are utilized; namely, binary phase-shift keying (BPSK), zero-mean white Gaussian noise with unit variance (WGN), and the first-order autoregressive signal (AR(1)) generated by $x(k) = 0.8x(k-1) + m(k)$, where $m(k)$ is a WGN signal. The signal-to-noise ratio (SNR) is adopted as 14 dB; i.e., the variance of additive noise signal is $\sigma_n^2 = 0.04$. The regularization factor and initialization vector are chosen as $\delta = 10^{-12}$ and $\mathbf{w}(0) = [0 \ \cdots \ 0]^T$, respectively. The simple choice constraint vector (CV) is adopted as $\boldsymbol{\gamma}(k)$ [12, 25]. The convergence factor of the AP algorithm is informed at each figure. The number of iterations is 10,000, and the learning curves and update rates are computed by averaging the results of 1000 independent trials.

6.1. Numerical examples

In this subsection, threshold parameter $\overline{\gamma}_1$ is estimated for the ST-SM-AP algorithm using the proposed approach in Section 4 (when $L = 2$) to obtain the desired results by chosen update rates $p = 0.15$, 0.30, and 0.44. The approximated threshold parameters for $L = 2$ as well as $p = 0.15$, 0.30, and 0.45 are listed in Table 1.

Table 1

Value of estimated $\overline{\gamma}_1$ for $p = 0.15$, 0.3, 0.44, and $L = 2$

p	0.15	0.3	0.44
$\overline{\gamma}$	0.343	0.2474	0.1903

Moreover, the update rates of the ST-SM-AP employing the estimated threshold parameters and 3 different input signals in 10,000 iterations are described in Table 2. We can observe that, by employing the computed $\overline{\gamma}_1$s, the resulting update rates listed in Table 2 are close to the values of p. Hence, using the estimating $\overline{\gamma}_1$s, the ST-SM-AP algorithm has censored non-informative suitably.

Table 2

Resulting update rates employing computed $\overline{\gamma}_1$ for $p = 0.15$, 0.3, 0.44, and $L = 2$ for ST-SM-AP algorithm and three different input signals

Input signal	p		
	0.15	0.3	0.44
BPSK	0.1498	0.2994	0.4392
WGN	0.1508	0.3007	0.4410
AR(1)	0.1514	0.3015	0.4417

Figures 2a–2i illustrate the MSE learning curves of the AP and ST-SM-AP algorithms (when $L = 2$) for input signals BPSK, WGN, and AR(1) and the three estimated $\overline{\gamma}_1$s for $p = 0.15$, 0.3, and 0.44.

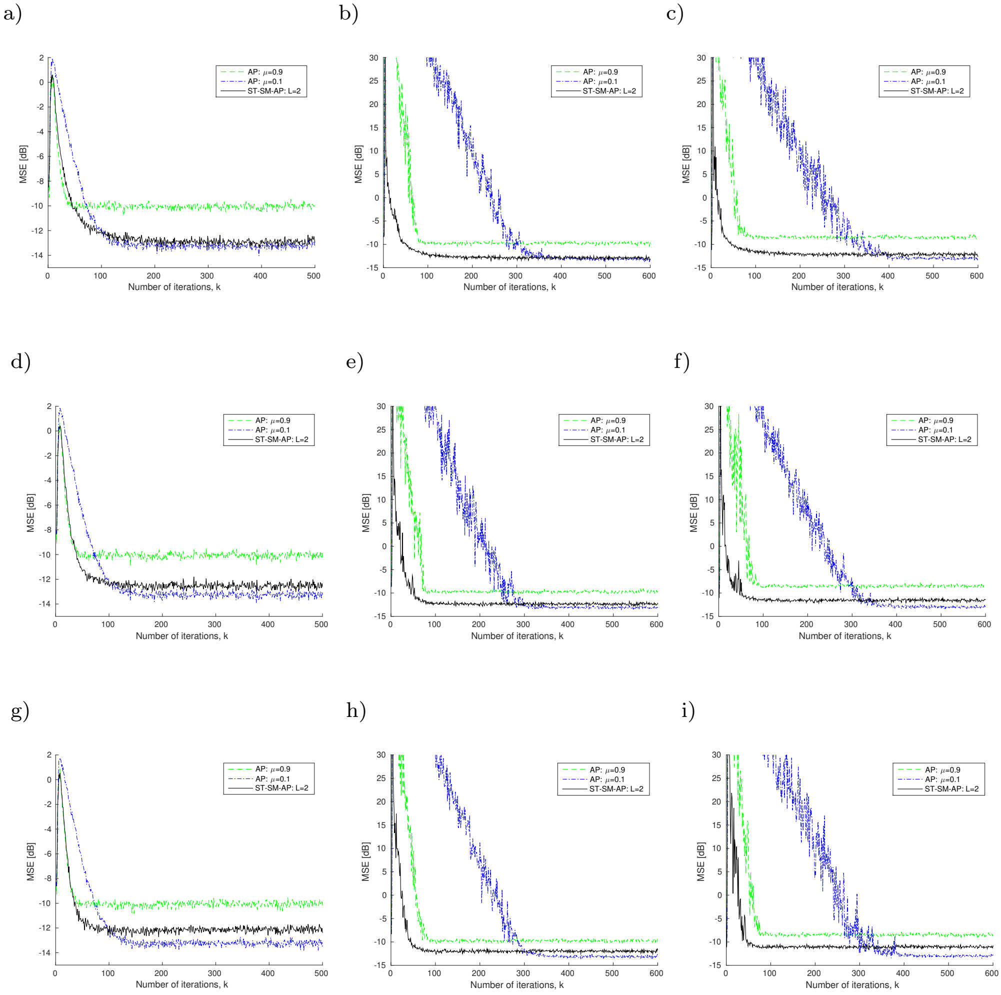

Figure 2. MSE learning curves of AP and ST-SM-AP algorithms for $L = 2$, considering: a) $\overline{\gamma}_1 = 0.343$ and BPSK input signal; b) $\overline{\gamma}_1 = 0.343$ and WGN input signal; c) $\overline{\gamma}_1 = 0.343$ and AR(1) input signal; d) $\overline{\gamma}_1 = 0.2474$ and BPSK input signal; e) $\overline{\gamma}_1 = 0.2474$ and WGN input signal; f) $\overline{\gamma}_1 = 0.2474$ and AR(1) input signal; g) $\overline{\gamma}_1 = 0.1903$ and BPSK input signal; h) $\overline{\gamma}_1 = 0.1903$ and WGN input signal; i) $\overline{\gamma}_1 = 0.1903$ and AR(1) input signal

Two different step-sizes for the AP algorithm (0.1 and 0.9) are selected. The update rates of the ST-SM-AP algorithm for these figures are presented in Table 2.

We can see that, when the step-size of the AP algorithm is large, the AP algorithm has a convergence rate as fast as the ST-SM-AP algorithm; however, the MSE of the ST-SM-AP algorithm is superior to that of the AP algorithm. When we adopt a small step-size for the AP algorithm to reach the MSE of the ST-SM-AP algorithm, the convergence velocity of the AP algorithm decreases remarkably. Also, note that the ST-SM-AP algorithm has an extremely lower computational cost as compared to the AP algorithm. Therefore, the ST-SM-AP algorithm can easily obtain a superior performance to the AP algorithm in big data applications.

To present an example using the DT-SM-AP algorithm, the AP, ST-SM-AP, and DT-SM-AP algorithms have been executed to identify $\mathbf{w}_o$ at the presence of an outlier signal. All of the parameters of the tested algorithms are chosen in an identical fashion to the previous example. The outlier added to the desired signal is generated by a Bernoulli process, which takes 1 with a probability of 0.05, then it is multiplied by uniformly distributed random numbers from interval $(0, 40)$.

To detect outliers and prevent updating adaptive coefficients for irrelevant data using the DT-SM-AP algorithm, the upper threshold is selected as $\overline{\gamma}_2 = 1$. In this case, the update rates of DT-SM-AP for $L = 2$ as well as the BPSK, WGN, and AR(1) input signals are presented in Table 3. As can be seen, the resulting update rates are close to the values of p in the presence of an outlier signal. Therefore, by adopting the estimated $\overline{\gamma}_1$s and $\overline{\gamma}_2$, the DT-SM-AP algorithm has censored the incoming data as we desired.

Table 3

Resulting update rates employing computed $\overline{\gamma}_1$ for $p = 0.15$, 0.3, 0.44, $L = 2$, and $\overline{\gamma}_2 = 1$ for DT-SM-AP algorithm and three different input signals

Input signal	p		
	0.15	0.3	0.44
BPSK	0.1496	0.2997	0.4395
WGN	0.1507	0.3010	0.4409
AR(1)	0.1516	0.3017	0.4418

Also, Figures 3a–3i show the misalignment curves of the AP, ST-SM-AP, and DT-SM-AP algorithms for $L = 2$ as well as the BPSK, WGN, and AR(1) input signals. Similar to the previous example, two step-sizes in the AP algorithm (0.1 and 0.9) are chosen such that the small μ results in a low MSE and low convergence speed; however, the large μ leads to a high MSE and high convergence velocity. Contrary to the update rates of the DT-SM-AP that is provided in Table 3, the update rates of the ST-SM-AP algorithm are greater than 80% due to the unnecessary updates in the presence of outliers. Therefore, the DT-SM-AP algorithm performs better in the presence of outliers when compared to the ST-SM-AP and AP algorithms.

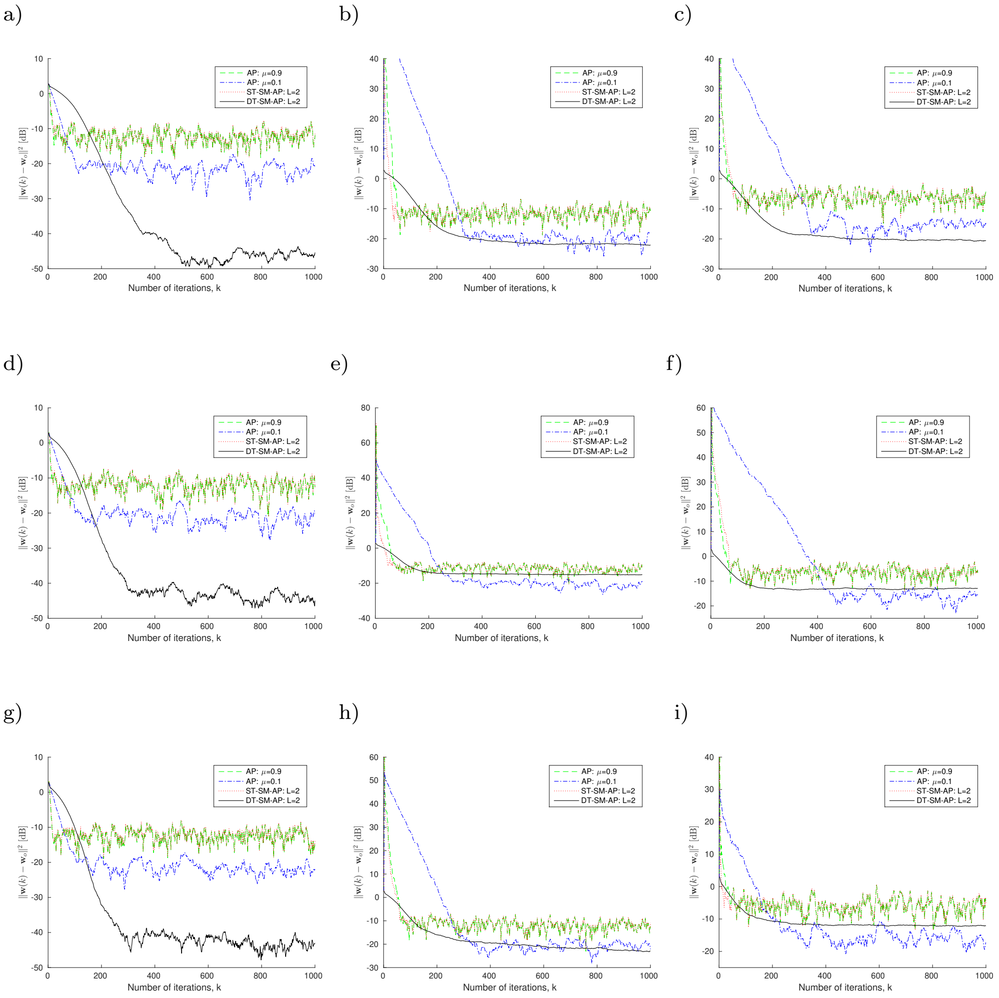

Figure 3. Misalignment curves of AP, ST-SM-AP, and DT-SM-AP algorithms for $L = 2$, considering: a) $\overline{\gamma}_1 = 0.343$ and BPSK input signal; b) $\overline{\gamma}_1 = 0.343$ and WGN input signal; c) $\overline{\gamma}_1 = 0.343$ and AR(1) input signal; d) $\overline{\gamma}_1 = 0.2474$ and BPSK input signal; e) $\overline{\gamma}_1 = 0.2474$ and WGN input signal; f) $\overline{\gamma}_1 = 0.2474$ and AR(1) input signal; g) $\overline{\gamma}_1 = 0.1903$ and BPSK input signal; h) $\overline{\gamma}_1 = 0.1903$ and WGN input signal; i) $\overline{\gamma}_1 = 0.1903$ and AR(1) input signal

6.2. Real-life example

In this subsection, we utilize the AP, ST-SM-AP, and DT-SM-AP algorithms to identify a measured unknown system corresponding to the room impulse response (RIR) tested in [16, 21]. For the AP algorithm, the step-size is chosen as $\mu = 0.9$. For the ST-SM-AP algorithm, $L = 1$ and $\overline{\gamma}_1 = 0.1692$. Also, for the DT-SM-AP algorithm,

$L = 1$, $\overline{\gamma}_1 = 0.1692$, and $\overline{\gamma}_2 = 1$. The SNR is adopted as 20 dB, and the input signal is a WGN signal. The initialization vector and regularization parameter are selected as the null vector and 10^{-12}, respectively.

Figures 4a and 4b present the MSE learning curves and the values of $E\|\mathbf{w}(k) - \mathbf{w}_*\|$, respectively, when the unknown system to be identified is the RIR. As can be seen, the ST-SM-AP and DT-SM-AP algorithms have superior performances to the AP algorithm. Moreover, the update rate of the ST-SM-AP and DT-SM-AP algorithms are 53.34% and 47.61%, respectively, while the AP algorithm updates the adaptive coefficients for all iterations. Therefore, besides the superior performance, the proposed algorithms require lower computational costs.

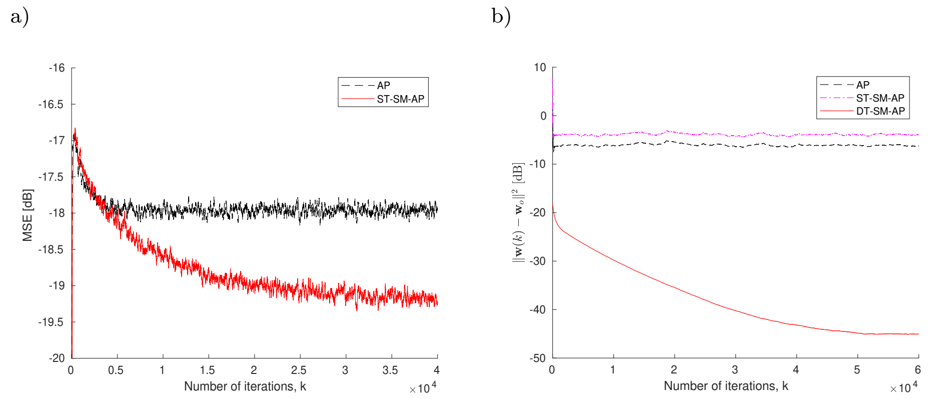

Figure 4. Results of proposed algorithms applied to identification of measured RIR: a) MSE learning curves of ST-SM-AP algorithms; b) value of $E\|\mathbf{w}(k) - \mathbf{w}_*\|$ for ST-SM-AP, DT-SM-AP, and AP algorithms

7. Conclusions

In this paper, we have proposed the single-threshold set-membership affine projection (ST-SM-AP) and the double-threshold set-membership affine projection (DT-SM-AP) algorithms to employ data abundance and censor non-informative and irrelevant data. To this end, the threshold parameter is approximated to obtain the desired update rate. Indeed, the probability distribution function of the additive noise signal and the excess mean-squared error in steady-state have been utilized to estimate the threshold parameter of the ST-SM-AP and DT-SM-AP algorithms. The ST-SM-AP algorithm prevents updating for non-informative data, whereas the DT-SM-AP algorithm avoids updating for non-informative and irrelevant data. The numerical results and real-life examples corroborate the superiority of the ST-SM-AP and DT-SM-AP algorithms when compared to the conventional affine projection algorithm.

References

[1] Apolinario J.A., de Campos M.L.R.: On efficient implementations of the set--membership NLMS algorithm for real-time applications. In: *2006 International Telecommunications Symposium* Fortaleza, Ceara, Brazil, pp. 275–278, 2006.

[2] Bhotto M.Z.A., Antoniou A.: A robust constrained set-membership affine-projection adaptive-filtering algorithm, *IEEE Transactions on Signal Processing*, vol. 60, pp. 73–81, 2012.

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

[4] Diniz P.S.R, Yazdanpanah H.: Improved set-membership partial-update affine projection algorithm. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016)*, Shanghai, China, pp. 4174–4178, 2016.

[5] Diniz P.S.R, Yazdanpanah H.: Data censoring with set-membership algorithms. In: *IEEE Global Conference on Signal and Information Processing (GlobalSIP 2017)*, Montreal, Canada, pp. 121–125, 2017.

[6] Diniz P.S.R.: *Adaptive Filtering: Algorithms and Practical Implementation*, 4th edition, New York, USA, Springer, 2013.

[7] Diniz P.S.R.: On Data-Selective Adaptive Filtering, *IEEE Transactions on Signal Processing*, vol. 66(16), pp. 4239–4252, 2018.

[8] Han S., De Maio S., Carotenuto V., Pallotta L., Huang X.: Censoring outliers in radar data: an approximate ML approach and its analysis, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55(2), pp. 534–546, 2019.

[9] Juneja A., Das N.N.: Big data quality framework: pre-processing data in weather monitoring application. In: *International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon 2019)*, Faridabad, India, pp. 559–563, 2019.

[10] Lima M.V.S, Diniz P.S.R.: Steady-state analysis of the set-membership affine projection algorithm. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2010)*, Dallas, USA, pp. 3802–3805, 2010.

[11] Lima M.V.S, Diniz P.S.R.: Steady-state MSE performance of the set-membership affine projection algorithm, *Circuits, Systems and Signal Processing*, vol. 32, pp. 1811–1837, 2013.

[12] Martins W.A., Lima M.V.S., Diniz P.S.R., Ferreira T.N.: Optimal constraint vectors for set-membership affine projection algorithms, *Signal Processing*, vol. 134, pp. 285–294, 2017.

[13] Meng F., Liu H., Shen X., et al.: Optimal prediction and update for box set--membership filter, *IEEE Access*, vol. 7, pp. 41636–41646, 2019.

[14] Msechu E.J., Giannakis G.B.: Decentralized data selection for MAP estimation: a censoring and quantization approach. In: *14th International Conference on Information Fusion*, Chicago, IL, USA, pp. 1–8, 2011.

[15] Msechu E.J., Giannakis G.B.: Sensor-centric data reduction for estimation with WSNs via censoring and quantization. *IEEE Transactions on Signal Processing*, vol. 60, pp. 400–414, 2012.

[16] Stewart R., Sandler M.: Database of omnidirectional and B-format room impulse responses. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2010)*, Dallas, USA, pp. 165–168, 2010.

[17] Takahashi N., Yamada I.: Steady-state mean-square performance analysis of a relaxed set-membership NLMS algorithm by the energy conservation argument, *IEEE Transactions on Signal Processing*, vol. 57, pp. 3361–3372, 2009.

[18] Wang Z., Shen X., Zhu Y., Pan J.: A tighter set-membership filter for some nonlinear dynamic systems, *IEEE Access*, vol. 6, pp. 25351–25362, 2018.

[19] Werner S., Diniz P.S.R.: Set-membership affine projection algorithm. *IEEE Signal Processing Letters*, vol. 8, pp. 231–235, 2001.

[20] Yazdanpanah H., Diniz P.S.R, Lima M.V.S.: A simple set-membership affine projection algorithm for sparse system modeling. In: *24th European Signal Processing Conference (EUSIPCO 2016)*, Budapest, Hungary, pp. 1798–1802, 2016.

[21] Yazdanpanah H., Diniz P.S.R, Lima M.V.S.: Low-complexity feature stochastic gradient algorithm for block-lowpass systems, *IEEE Access*, vol. 7, pp. 141587–141593, 2019.

[22] Yazdanpanah H., Diniz P.S.R. (2017) New trinion and quaternion set-membership affine projection algorithms. *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 64, pp. 216–220, 2017.

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

[24] Yazdanpanah H., Lima M.V.S, Diniz P.S.R.: 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, July 2016, pp. 1–5, 2016.

[25] Yazdanpanah H., Lima M.V.S, Diniz P.S.R.: On the robustness of set-membership adaptive filtering algorithms, *EURASIP Journal on Advances in Signal Processing*, vol. 72, pp. 1–12, 2017.

[26] Yazdanpanah H.: *On data-selective learning*, Federal University of Rio de Janeiro, 2018.

[27] Zardadi A.: Data selection with set-membership affine projection algorithm, *AIMS Electronics and Electrical Engineering*, vol. 3, pp. 359–369, 2019.

[28] Zhang S., Zhang J.: Set-membership NLMS algorithm with robust error bound. *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 61, pp. 536–540, 2014.

[29] Zheng Y., Niu R., Varshney P.K.: Sequential bayesian estimation with censored data for multi-sensor systems, *IEEE Transactions on Signal Processing*, vol. 62, pp. 2626–2641, 2014.

[30] Zhu H., Qian H., Luo X., Yang Y.: Adaptive queuing censoring for big data processing, *IEEE Signal Processing Letters*, vol. 25, pp. 610–614, 2018.

Affiliations

Gholamreza Karamali
Shahid Sattari Aeronautical University of Science and Technology, Faculty of Basic Sciences, South Mehrabad, Tehran, Iran, g_karamali@iust.ac.ir

Akram Zardadi
Payame Noor University (PNU), Department of Mathematics, P.O. Box, 19395-4697, Tehran, Iran, azardadi1990@yahoo.com

Hamid Reza Moradi
Shahid Sattari Aeronautical University of Science and Technology, Faculty of Basic Sciences, South Mehrabad, Tehran, Iran, hrmoradi@mshdiau.ac.ir

Received: 31.07.2019
Revised: 05.12.2019
Accepted: 09.12.2019