THE DATA-SELECTIVE CONSTRAINED AFFINE-PROJECTION ALGORITHM

Stefan Werner[1], *José A. Apolinário Jr.*[2], *Marcello L. R. de Campos*[3]

[1]Helsinki University of Technology
Signal Processing Laboratory
P.O.Box 3000
Espoo
FIN-02015 HUT, Finland
Stefan.Werner@hut.fi

[2]Instituto Militar de Engenharia
Depto. Engenharia Elétrica
Praça Gal. Tibúrcio, 80
Rio de Janeiro, RJ
22.290-270, Brazil
apolin@ieee.org

[3]Universidade Federal do Rio de Janeiro
COPPE/Prog. Engenharia Elétrica
P.O. Box 68504
Rio de Janeiro, RJ
21.945-970 Brazil
campos@lps.ufrj.br

ABSTRACT

This paper introduces a constrained version of the recently proposed set-membership affine projection algorithm based on the set-membership criteria for coefficient update. The algorithm is suitable for linearly-constrained minimum-variance filtering applications. The data selective property of the proposed algorithm greatly reduces the computational burden as compared with a nonselective approach. Simulation results show the good performance in terms convergence, final misadjustment, and reduced computational complexity.

1. INTRODUCTION

Adaptation algorithms which satisfy linear constraints encounter application in several areas of signal processing and communications, such as beamforming, spectral estimation, multiuser detection for communication systems, etc. A robust algorithm which does not require reinitialization and incorporates the constraints into the solution was first introduced by Frost [1]. More recently, other constrained adaptation algorithms were introduced which are tailored to specific applications or present advantageous performance regarding convergence and robustness (see, e.g., [2][3]).

The affine-projection (AP) algorithm is among the prominent unconstrained adaptation algorithms that may have a good compromise between fast convergence and low computational complexity. By adjusting the number of projections, performance can be controlled from that of the normalized least mean squares (NLMS) algorithm to that of the sliding-window recursive least squares (RLS) algorithm [4][5]. A constrained version of the affine-projection algorithm, the CAP algorithm, was proposed in [6] and was shown to achieve fast convergence. However, the fast convergence comes at the expense of a higher misadjustment.

In order to combat the conflicting requirements often encountered with most adaptive filtering algorithms, the objective function of the adaptive algorithm needs to be changed. In set-membership filtering (SMF) [7] an upper bound of the output estimation error is specified. The resulting adaptation algorithms are data-selective which in turn can reduce the computational complexity of the algorithms considerably. Furthermore, the sparse updating also results in a low misadjustment because the algorithms does not utilize the input data if it does not imply innovation. The set-membership affine projection (SM-AP) algorithm proposed in [8] generalized the work in [7] and [9], and was shown to achieve fast convergence and low misadjustment.

In this paper we apply the concept of set-membership filtering to the linearly constrained filtering problem in order to derive an efficient algorithm with low computational complexity and fast convergence. The new algorithm presented can be seen as a constrained version of the SM-AP algorithm and its recursions are similar to the constrained affine-projection (CAP) algorithm [6]. The new algorithm retain the fast convergence of the CAP algorithm, and low misadjustment is obtained due to the data selective property.

2. SET-MEMBERSHIP FILTERING

This section reviews the basic concepts of set-membership filtering (SMF). For a more detailed introduction to the concept of SMF, the reader is referred to [7]. In SMF, an upper bound is specified on the magnitude of the output estimation error $e_k = d_k - \mathbf{w}_k^T\mathbf{x}$. As a result of the bound constraint, the adaptive filtering algorithms derived within the framework of SMF will not perform filter update for all incoming signals, in other words they are data selective. In SMF, all vectors that belong to the *feasibility set*

$$\Theta = \bigcap_{(\mathbf{x},d)\in\mathcal{S}} \{\mathbf{w} \in \mathcal{R}^{N+1} : |d - \mathbf{w}^T\mathbf{x}| \leq \gamma\} \qquad (1)$$

are considered valid estimates, where $\mathcal{S}$ denotes the set of all possible input-desired data pairs $(\mathbf{x}, d)$ of interest. In many applications it is impossible to predict all possible data pairs and, therefore, adaptive methods work with the *membership* *sets* ψ_k constructed from the observed data pairs,

$$\psi_k = \bigcap_{i=1}^{k} \mathcal{H}_i \tag{2}$$

where

$$\mathcal{H}_k = \{\mathbf{w} \in \mathcal{R}^N : |d_k - \mathbf{w}^T\mathbf{x}_k| \leq \gamma\} \tag{3}$$

is the *constraint set* formed by the input data pair at time instant k. Note that the feasibility set is included in the membership set and if all possible data pairs are traversed up to time instant k, the membership becomes equal to the feasibility set. Since the membership is not easily solved for [7], simple adaptive approaches compute a point estimate provided part of the information in the membership set ψ_k, e.g., the information provided by the constraint set $\mathcal{H}_k$ like in the SM-NLMS [7] algorithm or by utilizing P past contraint sets like in the SM-AP algorithm [8].

3. SET-MEMBERSHIP CONSTRAINED AFFINE PROJECTION ALGORITHM

In linearly constrained adaptive filtering, the constraints are given by the set of equations

$$\mathbf{C}^T\mathbf{w} = \mathbf{f} \tag{4}$$

where $\mathbf{C}$ is a $(N+1)\times J$ constraint matrix and $\mathbf{f}$ is the vector of J constraint values. In our SMF formulation we want to design our filter such that the magnitude of estimation error is bounded. For this formulation we partition the membership set as $\psi_k = \psi_k^{k-P} \cap \psi_k^P$ where ψ_k^P corresponds to the intersection of the P past constraint sets, i.e.,

$$\psi_k^P = \bigcap_{i=k-P+1}^{k} \mathcal{H}_i \tag{5}$$

Next we consider the derivation of a data-selective algorithm whose coefficients belong to the hyperplane defined by equation (4) and also to the partitioned membership set, i.e., $\mathbf{C}^T\mathbf{w}_{k+1} = \mathbf{f}$ and $\mathbf{w}_{k+1} \in \psi_k^P$. Let us state the following optimization criterion whenever $\mathbf{w}_k \notin \psi_k^P$.

$$\begin{aligned} \mathbf{w}_{k+1} &= \arg\min \|\mathbf{w}_{k+1} - \mathbf{w}_k\|^2 \ \text{ subject to} \\ &\mathbf{C}^{\mathrm{T}}\mathbf{w}_{k+1} = \mathbf{f} \\ &\mathbf{d}_k - \mathbf{X}_k^{\mathrm{T}}\mathbf{w}_{k+1} = \mathbf{g}_k \end{aligned} \tag{6}$$

where

$$\begin{aligned} \mathbf{g}_k &= [g_k \; g_{k-1} \; \ldots \; g_{k-P+1}]^{\mathrm{T}} \\ \mathbf{d}_k &= [d_k \; d_{k-1} \; \ldots \; d_{k-P+1}]^{\mathrm{T}} \\ \mathbf{X}_k &= [\mathbf{x}_k \; \mathbf{x}_{k-1} \; \cdots \; \mathbf{x}_{k-P+1}] \end{aligned} \tag{7}$$

with $\mathbf{x}_k = [x_k \; x_{k-1} \; \ldots \; x_{k-N}]^{\mathrm{T}}$ and N the filter order. In order to guarantee that $\mathbf{w}_{k+1} \in \psi_k^P$ the elements of $\mathbf{g}_k$ are chosen such that $|g_{k-i+1}| \leq \gamma$ for $i = 1 \ldots P$. In the end of this section we consider a particular choice of the parameters g_{k-i+1} leading to a simplified algorithm. Using the method of Lagrange multipliers, the unconstrained function to be minimized may be expressed by

$$\begin{aligned} f(\mathbf{w}_{k+1}) = \|\mathbf{w}_{k+1} - \mathbf{w}_k\|^2 + \boldsymbol{\lambda}_1^{\mathrm{T}}[\mathbf{f} - \mathbf{C}^{\mathrm{T}}\mathbf{w}_{k+1}] \\ + \boldsymbol{\lambda}_2^{\mathrm{T}}[\mathbf{d}_k - \mathbf{X}_k^{\mathrm{T}}\mathbf{w}_{k+1} - \mathbf{g}_k] \end{aligned} \tag{8}$$

Setting the gradient of $f(\mathbf{w}_{k+1}, \boldsymbol{\lambda})$ with respect to $\mathbf{w}_{k+1}$ equal to zero yields

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mathbf{C}\frac{\boldsymbol{\lambda}_1}{2} + \mathbf{X}_k\frac{\boldsymbol{\lambda}_2}{2} \tag{9}$$

Solving for the constraints we get

$$\mathbf{w}_{k+1} = \mathbf{P}\left[\mathbf{w}_k + \mathbf{X}_k\mathbf{t}_k\right] + \mathbf{F} \tag{10}$$

where

$$\begin{aligned} \mathbf{t}_k &= \left(\mathbf{X}_k^{\mathrm{T}}\mathbf{P}\mathbf{X}_k\right)^{-1}\left(\mathbf{d}_k - \mathbf{X}_k^{\mathrm{T}}\mathbf{w}_k - \mathbf{g}_k\right) \\ &= \left(\mathbf{X}_k^{\mathrm{T}}\mathbf{P}\mathbf{X}_k\right)^{-1}\left(\mathbf{e}_k - \mathbf{g}_k\right) \end{aligned} \tag{11}$$

$$\mathbf{e}_k = [e_k \; \epsilon_{k-1} \; \ldots \; \epsilon_{k-P+1}]^{\mathrm{T}} \tag{12}$$

with $\epsilon_{k-i} = d_{k-i} - \mathbf{x}_{k-i}^T\mathbf{w}_k$ denoting the *a posteriori* error at iteration $k-i$. The matrix

$$\mathbf{P} = \mathbf{I} - \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T \tag{13}$$

is a projection matrix for a projection onto the homogeneous hyperplane defined by $\mathbf{C}^T\mathbf{w}_k = \mathbf{0}$, and the vector

$$\mathbf{F} = \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{f} \tag{14}$$

is used to move the projected solution back to the constraint hyperplane.

Now, let us look more closely on the constraint vector $\mathbf{g}_k$. Due to the data reusing property of the above algorithm we have $\mathbf{w}_k \in \mathcal{H}_{k-i+1}$, i.e., $|\epsilon_{k-i+1}| \leq \gamma$, for $i \neq 1$. Therefore, choosing $g_{k-i+1} = \epsilon_{k-i+1}$, for $i \neq 1$, will cancel all but the first element in the term $\mathbf{e}_k - \mathbf{g}_k$ of (11).

In the same way as with the SM-NLMS and the SM-AP algorithms we can choose g_k such that the *a posteriori* error lies on the closest boundary of $\mathcal{H}_k$, i.e., $g_k = \gamma \text{sign}(e_k)$. With the above choices we get

$$\mathbf{t}_k = \left(\mathbf{X}_k^{\mathrm{T}}\mathbf{P}\mathbf{X}_k\right)^{-1}\alpha_k e_k \mathbf{u}_1 \tag{15}$$

where $\mathbf{u}_1 = [1\;0\;\ldots\;0]^T$ and

$$\alpha_k = \begin{cases} 1 - \gamma/|e_k| & \text{if } |e_k| > \gamma \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

is the data dependent step-size. Note that for time instants $k < P$ only knowledge of $\mathcal{H}_i$ for $i = 1, \ldots, k$ can be assumed. If an update is needed for the initial time instants $k < P$, the algorithm is used with the k available constraint sets. The equations of the SM-CAP algorithm is summarized in Table 1, where a small constant δ was used to avoid the inversion of a possible null matrix. For comparison, Table 2 shows the CAP algorithm proposed in [6]. In both algorithms given below the simplification $\mathbf{P}\mathbf{w}_k + \mathbf{F} = \mathbf{w}_k$ should be avoided to prevent the solution to drift away from the constraint plane in a finite precision implementation [1].

Table 1. The set-membership constrained affine projection algorithm.

SM-CAP Algorithm
for each k
$\{$
$e_k = d_k - \mathbf{x}_k^T \mathbf{w}_k$
if $\lvert e_k \rvert > \gamma$
$\alpha_k = 1 - \gamma / \lvert e_k \rvert$
$\mathbf{t}_k = \left[\mathbf{X}_k^T \, \mathbf{P} \, \mathbf{X}_k + \delta \mathbf{I}\right]^{-1} \alpha_k e_k \mathbf{u}_1$
$\mathbf{w}_{k+1} = \mathbf{P}\left[\mathbf{w}_k + \mathbf{X}_k \, \mathbf{t}_k\right] + \mathbf{F}$
else
$\mathbf{w}_{k+1} = \mathbf{w}_k$
$\}$

Table 2. The constrained affine projection algorithm.

CAP Algorithm
for each k
$\{$
$\mathbf{e}_k = \mathbf{d}_k - \mathbf{X}_k^T \mathbf{w}_k$
$\mathbf{t}_k = \left[\mathbf{X}_k^T \, \mathbf{P} \, \mathbf{X}_k + \delta \mathbf{I}\right]^{-1} \mathbf{e}_k$
$\mathbf{w}_{k+1} = \mathbf{P}\left[\mathbf{w}_k + \mu \mathbf{X}_k \, \mathbf{t}_k\right] + \mathbf{F}$
$\}$

4. SIMULATIONS

4.1. Experiment 1

A first experiment was carried out in a system-identification problem where the filter coefficients were constrained to preserve linear phase at every iteration. For this example we made $N = 10$ and, in order to fulfill the linear phase requirement, we made

$$\mathbf{C} = \begin{bmatrix} \mathbf{I}_{N/2} \\ \mathbf{0}^T \\ -\mathbf{J}_{N/2} \end{bmatrix} \tag{17}$$

with $\mathbf{J}$ being a reversal matrix (an identity matrix with all lines in reversed order), and

$$\mathbf{f} = [0 \cdots 0]^T \tag{18}$$

This didactic setup was employed to show the improvement of the convergence speed when P is increased. Due to the symmetry of $\mathbf{C}$ and the fact that $\mathbf{f}$ is a null vector, more efficient structures could be used [10]. The input signal consists of zero-mean unity-variance colored noise with eigenvalue spread around 2000 and the reference signal was obtained after filtering the input by a linear-phase FIR filter and adding observation noise with variance equal to 10^{-10}. The value of γ in the SM-CAP algorithm was chosen equal to $3\sigma_n$. A higher value would result in less frequent updates but in slightly higher final misadjustment. Fig. 1 shows the learning curves for the SM-CAP and CAP algorithms for $P = 1$, $P = 2$, and $P = 4$. Fig. 1 clearly shows the increase in convergence speed obtained by increasing P as can be expected. It is also clear from this figure that the misadjustment with the SM-CAP algorithm is lower than the CAP algorithm, and that the misadjustment increases more slowly when P is increased. The only way for the CAP algorithm to achieve the low misadjustment of the SM-CAP is through the introduction of a step-size resulting in a slower convergence. Furthermore, in 500 iterations the SM-CAP algorithm performed updates in 485, 111, and 100 time instants for $P = 1$, $P = 2$, and $P = 4$, respectively. In other words, the SM-CAP algorithm with $P = 4$ had a better performance than the CAP algorithm while performing updates for only a fraction of data.

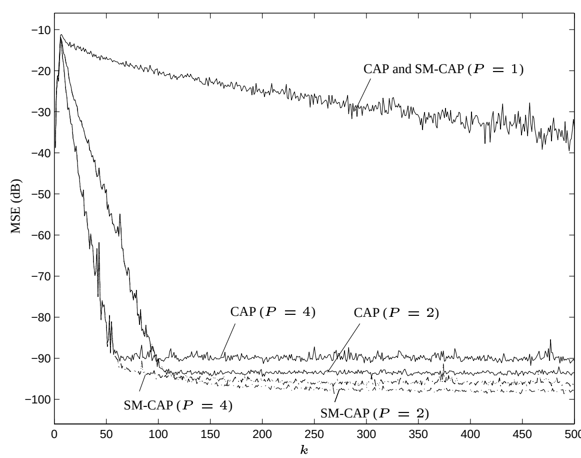

Fig. 1. Learning curves for the CAP and the SM-CAP algorithms with $P = 1$, $P = 2$, and $P = 4$ data reuses, $\sigma_n^2 = 10^{-10}$, $\gamma = 3\sigma_n$, and colored input signal.

4.2. Experiment 2

A second experiment was conducted where the received signal consists of three sinusoids in white noise:

$$x_k = \sin(0.3k\pi) + \sin(0.325k\pi) + \sin(0.7k\pi) + n_k \quad (19)$$

where n_k is white noise with power such that the SNR is 40dB. The filter is constrained to pass frequency component of 0.1rad/s and 0.25rad/s undistorted resulting in the following constraint matrix and vector:

$$\mathbf{C}^T = \begin{bmatrix} 1 & \cos(0.2\pi) & \cdots & \cos\left[(N-1)0.2\pi\right] \\ 1 & \cos(0.5\pi) & \cdots & \cos\left[(N-1)0.5\pi\right] \\ 0 & \sin(0.2\pi) & \cdots & \sin\left[(N-1)0.2\pi\right] \\ 0 & \sin(0.5\pi) & \cdots & \sin\left[(N-1)0.5\pi\right] \end{bmatrix} \quad (20)$$

$$\mathbf{f}^T = [1\ 1\ 0\ 0] \quad (21)$$

In this example the reference signal is set to zero, i.e., $e_k = -\mathbf{x}_k^T\mathbf{w}_k$. The norm of the mean output energy (MOE) is shown in Figure 2 for the SM-CAP and the CAP algorithms for $P = 3$. The threshold γ was set to $4\sigma_n$. A step size $\mu_{CAP} = 0.15$ was used with the CAP to obtain a steady-state close to the SM-CAP algorithm. We see from the figure that the SM-CAP curve is less noisy than the CAP curve during the initial 1500 iteratians. After the convergence both algorithm have similar steady-state value. In 5000 iterations, the average number of updates for the SM-CAP algorithm was 790 as compared with 5000 updates for the CAP algorithm.

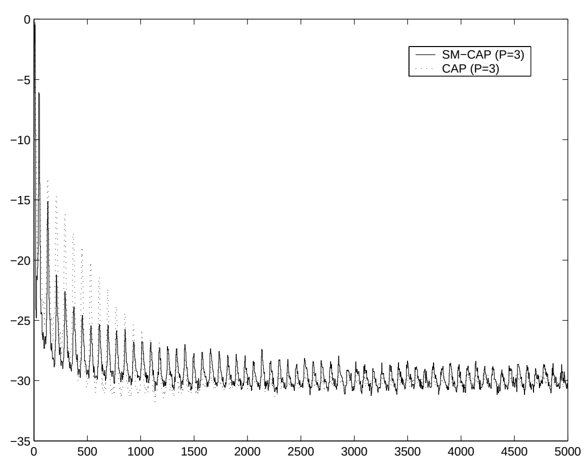

Fig. 2. The mean output power

5. CONCLUSIONS

A constrained version of the set-membership affine projection (SM-CAP) algorithm was presented. The algorithm is based on the concept of set-membership filtering, and utilizes consecutive data-pairs in order to construct a space of feasible solutions for the updates. The data selective feature can, in certain applications, reduce substantially the number of required updates as compared with the conventional constrained affine projection algorithm (CAP). Simulations confirmed that the proposed algorithm leads to fast convergence speed, low misadjustment, and a substantial reduction in the number of updates.

6. REFERENCES

[1] O. L. Frost III, "An algorithm for linearly constrained adaptive array processing," *Proceedings of the IEEE*, vol. 60, pp. 926–935, Aug. 1972.

[2] L. S. Resende, J. M. T. Romano, and M. G. Bellanger, "A fast least-squares algorithm for linearly constrained adaptive filtering," *IEEE Trans. Signal Processing*, vol. 44, pp. 1168–1174, May 1996.

[3] M. L. R. de Campos, S. Werner, and J. A. Apolinário Jr., "Householder-transform constrained LMS algorithms with reduced-rank updating," in *Proc. International Conference of Acoustics, Speech, and Signal Processing*, Phoenix, USA, pp. 1857–1960, 1999.

[4] M. L. R. de Campos, P. S. R. Diniz, and J. A Apolinário Jr., "On normalized data-reusing and affine-projections algorithms," in *Proc. IEEE International Conference on Electronics, Circuits, and Systems*, Pafos, Cyprus, 1999.

[5] M. Montazeri and P. Duhamel, "A set of algorithms linking NLMS and block RLS algorithms," *IEEE Trans. Signal Processing*, vol. 43, pp. 444–453, Feb. 1995.

[6] M. L.R. de Campos, J. A Apolinário Jr., and P. S. R. Diniz, "The constrained affine projection algorithm – development and convergence issues," in *Proc. First Balkan Conference on Signal Processing, Communications, Circuits and Systems*, Istanbul, Turkey, 2000.

[7] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y.-F. Huang, "Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size," *IEEE Signal Processing Letters*, vol. 5, pp. 111-114, May 1998.

[8] S. Werner and P. S. R. Diniz, "Set-membership affine projection algorithm," submitted for publication to IEEE Signal Processing Letters, 2000.

[9] P. S. R. Diniz and S. Werner, "Set-membership binormalized data-reusing LMS algorithms," *IFAC Symposium on System Identification*, Santa Barbara, USA, 2000.

[10] L. S. Resende, J. M. T. Romano, and M. G. Bellanger, "Simplified FLS algorithm for linear phase adaptive filtering," in *Proc. European Signal Processing Conference*, vol. III, Rodes, Greece, pp. 1237–1240, 1998.