HOUSEHOLDER-TRANSFORM CONSTRAINED LMS ALGORITHMS WITH REDUCED-RANK UPDATING

M. L. R. de Campos,[1,2] *S. Werner,*[2] *and J. A. Apolinário Jr.*[3]

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

[2] Helsinki University of Technology
Laboratory of Telecommunications Technology
P.O.Box 3000
FIN-02015 HUT, Finland
Stefan.Werner@hut.fi

[3] Instituto Militar de Engenharia
Depto. de Engenharia Elétrica
Praça Gal. Tibúrcio, 80
Praia Vermelha, Rio de Janeiro, RJ
22.290-270, Brazil
apolin@aquarius.ime.eb.br

ABSTRACT

This paper proposes a new approach to linearly-constrained adaptive filtering, where successive Householder transformations are incorporated in the algorithm update equation in order to reduce computational complexity and coefficient-error norm. We show the derivation of two new algorithms, namely the unnormalized and the normalized Householder-transform constrained LMS algorithms (HCLMS and NHCLMS, respectively). Although the derivation is carried out based on the constrained LMS (CLMS) algorithm, the technique can be applied to other constrained algorithms as well. Simulation results of a linearly-constrained minimum-variance problem show that in finite-precision implementation the coefficient-error norms obtained with the new algorithms are smaller than those obtained with the CLMS and the normalized CLMS algorithms.

1. INTRODUCTION

Linearly-constrained adaptive filters have been used in many applications, such as antenna arrays and interference suppression in CDMA. In general, constrained adaptation algorithms are derived from their unconstrained versions, e.g., the constrained LMS (CLMS) algorithm [1], the fast constrained RLS algorithm [2], the constrained normalized and binormalized LMS algorithms [3], and the constrained quasi-Newton (QN) algorithm [4]. In all these algorithms the direction of updating is premultiplied by a rank-deficient projection matrix, which renders them not optimal in the sense of computational complexity and finite-precision arithmetic behavior. Furthermore, these algorithms have in common a correction factor to ensure the constraints at every iteration. If this correction factor is not applied, then perturbations due to roundoff errors in certain directions cannot be suppressed and coefficient divergence will occur [5].

In this paper we address these problems by suitably transforming the input-signal vector so that the algorithm operates on a reduced-dimension subspace and therefore does not require updating of all its coefficients. Furthermore, the subspace where the coefficient updating is performed is orthogonal to the subspace spanned by the constraint matrix. No correction terms need be applied and the solution satisfies the constraints exactly at every iteration. The proposed algorithm has lower computational complexity than the CLMS algorithm and also smaller coefficient-error norm when implemented in finite precision. In addition, the more constraints introduced, the more economical is the algorithm and the smaller is the coefficient-error norm compared to the CLMS algorithm. When compared to the generalized side-lobe canceler (GSC) a significant reduction in computational complexity is achieved for applications with more than one constraint [5].

For the purpose of comparison, in this paper we show a finite-precision implementation of a minimum-variance constrained filter, where reducing computational complexity and satisfying constraints exactly is of crucial importance. We anticipate that these advantages will be equally evident in several other applications such as antenna arrays, mobile communications, and others.

2. OPTIMAL LINEARLY-CONSTRAINED MINIMUM-VARIANCE FILTER

The optimal (LCMV) filter in the sense of the minimum mean squared error (MSE) is the one that minimizes the following objective function:

$$\xi_w = \frac{1}{2}||\mathbf{R}^{1/2}\mathbf{w}||^2 \tag{1}$$

subjected to the set of linear constraints defined by

$$\mathbf{C}^T\mathbf{w} = \mathbf{f} \tag{2}$$

where $\mathbf{w}$ is a vector of coefficients of length M, $\mathbf{R}^{1/2}$ is the square-root factor of the the autocorrelation matrix of the input signal, $\mathbf{C}$ is the $M \times p$ constraint matrix, and $\mathbf{f}$ is the $p \times 1$ gain vector. The reference signal in this case has been chosen equal to zero without loss of generality.

By using the method of Lagrange multipliers, the constrained optimization problem becomes an unconstrained optimization problem in which the objective function minimized is

$$\xi_w = \frac{1}{2}||\mathbf{R}^{1/2}\mathbf{w}||^2 + \boldsymbol{\lambda}^T \left(\mathbf{C}^T\mathbf{w} - \mathbf{f}\right) \tag{3}$$

with $\boldsymbol{\lambda} \in \mathbb{R}^{p\times 1}$. The gradient of (3) with respect to the coefficients is

$$\nabla_\xi = \mathbf{R}\mathbf{w} + \mathbf{C}\boldsymbol{\lambda} \tag{4}$$

The optimal coefficients are the ones for which $\nabla_\xi = 0$, i.e.,

$$\mathbf{w}_{opt} = -\mathbf{R}^{-1}\mathbf{C}\boldsymbol{\lambda} \tag{5}$$

with

$$\boldsymbol{\lambda} = -(\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C})^{-1}\mathbf{f} \tag{6}$$

that exists if $(\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C})^{-1}$ exists. Therefore,

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{C}(\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C})^{-1}\mathbf{f} \tag{7}$$

The equation above bears the difficulty of knowing in real-time the inverse of the input-signal autocorrelation matrix, $\mathbf{R}^{-1}$. A much more practical approach is to produce an estimate of $\mathbf{w}_{opt}$ recursively at every iteration. As time proceeds, the estimate is improved such that convergence in the mean to the optimal solution may eventually be achieved. Frost [1] has proposed an algorithm to estimate $\mathbf{w}_{opt}$ based on the gradient method or, more specifically, based on the LMS algorithm for adaptive filtering.

2.1. The Constrained LMS Algorithm

Let $\mathbf{w}_k$ denote the estimate of $\mathbf{w}_{opt}$ at time instant k. Then, according to the LMS algorithm, $\mathbf{w}_{k+1}$ is given by

$$\begin{aligned}\mathbf{w}_{k+1} &= \mathbf{w}_k - \mu\nabla_{\xi,k} \\ &= \mathbf{w}_k - \mu\left[\mathbf{R}_k\mathbf{w}_k + \mathbf{C}\boldsymbol{\lambda}_k\right]\end{aligned} \tag{8}$$

where $\nabla_{\xi,k}$ is the estimate of ∇_ξ at time instant k and μ is a step-size used to control speed of convergence and steady-state estimation error.

As the constraints must be satisfied at every iteration,

$$\mathbf{C}^T\mathbf{w}_{k+1} = \mathbf{f} \tag{9}$$

Therefore,

$$\mathbf{f} = \mathbf{C}^T\mathbf{w}_k - \mu\mathbf{C}^T\left[\mathbf{R}_k\mathbf{w}_k + \mathbf{C}\boldsymbol{\lambda}_k\right] \tag{10}$$

Solving for $\boldsymbol{\lambda}_k$ gives

$$\boldsymbol{\lambda}_k = (\mathbf{C}^T\mathbf{C})^{-1}\left[\frac{\mathbf{C}^T\mathbf{w}_k - \mathbf{f}}{\mu} - \mathbf{C}^T\mathbf{R}_k\mathbf{w}_k\right] \tag{11}$$

which yields a unique solution if $(\mathbf{C}^T\mathbf{C})^{-1}$ exists. The algorithm uses as an estimate of the input-signal autocorrelation matrix $\mathbf{R}_k$ the outer product of the input-signal vector by itself, i.e., $\mathbf{R}_k = \mathbf{x}_k\mathbf{x}_k^T$. In this case, the coefficient-update equation is as follows:

$$\begin{aligned}\mathbf{w}_{k+1} &= \mathbf{w}_k - \mu y_k\left[\mathbf{I} - \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\right]\mathbf{x}_k \\ &\quad + \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}(\mathbf{f} - \mathbf{C}^T\mathbf{w}_k) \\ &= \mathbf{w}_k - \mu y_k\mathbf{P}_\perp\mathbf{x}_k \\ &\quad + \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}(\mathbf{f} - \mathbf{C}^T\mathbf{w}_k)\end{aligned} \tag{12}$$

where

$$y_k = \mathbf{w}_k^T\mathbf{x}_k \tag{13}$$

is the filter output, equal to the output error in this case of zero reference signal, $\mathbf{I}$ is the Mth-order identity matrix, and

$$\mathbf{P}_\perp = \mathbf{I} - \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T \tag{14}$$

is the projection matrix onto the subspace orthogonal to the subspace spanned by the constraint matrix.

A normalized version of the CLMS algorithm, namely the NCLMS algorithm, can be easily derived [3]; the update equation becomes:

$$\begin{aligned}\mathbf{w}_{k+1} &= \mathbf{w}_k - \mu\frac{y_k}{\mathbf{x}_k^T\mathbf{P}_\perp\mathbf{x}_k}\mathbf{P}_\perp\mathbf{x}_k \\ &\quad + \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}(\mathbf{f} - \mathbf{C}^T\mathbf{w}_k)\end{aligned} \tag{15}$$

The necessity of the last term in (12) and (15) may be surprising, for it is expected that all $\mathbf{w}_k$ satisfy the constraint and, therefore, this last term should be equal to zero. In practice, however, this term shall be included to prevent divergence of the coefficients in a limited-precision arithmetic machine [5]. Without this term, the updating of the coefficient vector in the CLMS algorithm is carried out in a direction given by the projection of the input-signal vector onto the subspace orthogonal to the subspace spanned by $\mathbf{C}$. Therefore, any direction within this subspace never contributes to correct the coefficient vector. Furthermore, perturbations introduced in this direction cannot be corrected by the adaptation of the algorithm and may cause a cumulative error effect [5] in the coefficient estimates. The importance of the last term in (12) is related to the correction of these perturbations introduced in the coefficient-update equation in the direction not excited by vector $\mathbf{P}_\perp\mathbf{x}_k$. The same reasoning can be applied to the RLS algorithm presented in [2] and to the QN algorithm presented in [4].

3. HOUSEHOLDER-TRANSFORM CONSTRAINED LMS (HCLMS) ALGORITHM

In the updating of the coefficients, the use of a projected input-signal vector which is nonpersistently exciting in a known direction is obviously suboptimal and dangerous. Besides the lack of capacity to correct perturbations in a direction not excited, an extra term of magnitude proportional to the perturbation must be calculated adding to algorithm complexity.

The HCLMS algorithm to be derived in this section performs a rotation on vector $\mathbf{P}_\perp \mathbf{x}_k$ in order to make sure that $\mathbf{w}_k$ is never perturbed in a direction not excited by $\mathbf{P}_\perp \mathbf{x}_k$. This can be done if an orthogonal rotation matrix $\mathbf{Q}$ is used as the transformation that will generate a modified coefficient vector $\bar{\mathbf{w}}_k$ that relates to $\mathbf{w}_k$ according to

$$\bar{\mathbf{w}}_k = \mathbf{Q}\mathbf{w}_k \tag{16}$$

If we choose the matrix $\mathbf{Q}$ such that $\mathbf{Q}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{Q} = \mathbf{I}$ and

$$\bar{\mathbf{C}}(\bar{\mathbf{C}}^T\bar{\mathbf{C}})^{-1}\bar{\mathbf{C}}^T = \begin{bmatrix} \mathbf{I}_{p\times p} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{17}$$

then $\bar{\mathbf{C}} = \mathbf{Q}\mathbf{C}$ satisfies $\mathbf{f} = \bar{\mathbf{C}}^T\bar{\mathbf{w}}_{k+1}$ and the transformed projection matrix is such that

$$\begin{aligned}\bar{\mathbf{P}}_\perp &= \mathbf{Q}\mathbf{P}_\perp\mathbf{Q}^T \\ &= \begin{bmatrix} \mathbf{0}_{p\times p} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}\end{aligned} \tag{18}$$

If the first p elements of $\bar{\mathbf{w}}_0$, $\hat{\mathbf{w}}_0$, are equal to the first p elements of vector $\bar{\mathbf{C}}(\bar{\mathbf{C}}^T\bar{\mathbf{C}})^{-1}\mathbf{f}$, then they need not be updated. The update equation of the HCLMS algorithm becomes

$$\bar{\mathbf{w}}_{k+1} = \begin{bmatrix} \hat{\mathbf{w}}_0 \\ \breve{\mathbf{w}}_{k+1} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{w}}_0 \\ \breve{\mathbf{w}}_k \end{bmatrix} - \mu y_k \begin{bmatrix} \mathbf{0} \\ \breve{\mathbf{x}}_k \end{bmatrix} \tag{19}$$

where $\breve{\mathbf{w}}_k$ and $\breve{\mathbf{x}}_k$ denote the $M-p$ last elements of vectors $\bar{\mathbf{w}}_k$ and $\bar{\mathbf{x}}_k$, respectively. Note that vector $\bar{\mathbf{C}}(\bar{\mathbf{C}}^T\bar{\mathbf{C}})^{-1}\mathbf{f}$ has only p nonzero elements.

Although the solution $\bar{\mathbf{w}}_k$ is biased by a transformation $\mathbf{Q}$, the output signal and, consequently, the output error are not modified by the transformation. We conclude, therefore, that the HCLMS algorithm minimizes the same objective function minimized by the CLMS algorithm.

Matrix $\mathbf{Q}$ may be constructed with successive Householder transformations [6] applied onto each of the p columns of matrix $\mathbf{CL}$, where $\mathbf{L}$ is the square-root factor of matrix $(\mathbf{C}^T\mathbf{C})^{-1}$, i.e., $\mathbf{L}\mathbf{L}^T = (\mathbf{C}^T\mathbf{C})^{-1}$. Let $\mathbf{Q}_i$ be the Householder transformation that produces zero entries in the last $M-i$ positions of the ith column of matrix $\mathbf{CL}$, then

$$\mathbf{Q} = \mathbf{Q}_p \cdots \mathbf{Q}_2\mathbf{Q}_1 \tag{20}$$

where

$$\mathbf{Q}_i = \begin{bmatrix} \mathbf{I}_{i-1\times i-1} & \mathbf{0}^T \\ \mathbf{0} & \bar{\mathbf{Q}}_i \end{bmatrix} \tag{21}$$

and matrix $\bar{\mathbf{Q}}_i$ is an ordinary $(M-i+1)\times(M-i+1)$ Householder transformation matrix.

Matrix $\mathbf{CL}$ has orthonormal columns, which means that matrix $\mathbf{QCL}$ is upper diagonal with ± 1 entries. Therefore, Eqs. (17) and (18) are satisfied and the algorithm can update the coefficients in a subspace with reduced dimension. The entries of vector $\mathbf{w}_k$ which lie in the subspace of the constraints need not be updated.

Likewise the CLMS algorithm, a normalized version of the HCLMS algorithm, namely the NHCLMS algorithm, can be easily derived and its update equation is

$$\bar{\mathbf{w}}_{k+1} = \begin{bmatrix} \hat{\mathbf{w}}_0 \\ \breve{\mathbf{w}}_{k+1} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{w}}_0 \\ \breve{\mathbf{w}}_k \end{bmatrix} - \mu \frac{y_k}{\breve{\mathbf{x}}_k^T\breve{\mathbf{x}}_k} \begin{bmatrix} \mathbf{0} \\ \breve{\mathbf{x}}_k \end{bmatrix} \tag{22}$$

4. COMPUTATIONAL COMPLEXITY

Although the HCLMS algorithm presented is indeed a transform-domain algorithm, the transformation performed onto the input-signal vector can be very efficiently coded and only the last $M-p$ elements of vector $\bar{\mathbf{x}}_k$ need be calculated and only the last $M-p$ elements of the coefficient vector need be updated. Table 1 shows the number of additions, multiplications, and divisions per iteration necessary for the HCLMS algorithm, the CLMS algorithm, and their normalized versions. For all algorithms the operation count was done by considering their most efficient implementation.

Table 1: Computational Complexity

ALG.	ADD.	MULT.	DIV.
CLMS	$(2p+2)M$ $-(p+1)$	$(2p+2)M+1$	0
NCLMS	$(3p+3)M$ $-(p+2)$	$(3p+3)M+1$	1
HCLMS	$(2p+2)M$ $-(p^2+p+1)$	$(2p+2)M$ $-(p^2-1)$	0
NHCLMS	$(2p+3)M$ $-(p^2+2p+2)$	$(2P+3)M$ $-(p^2+p-1)$	1

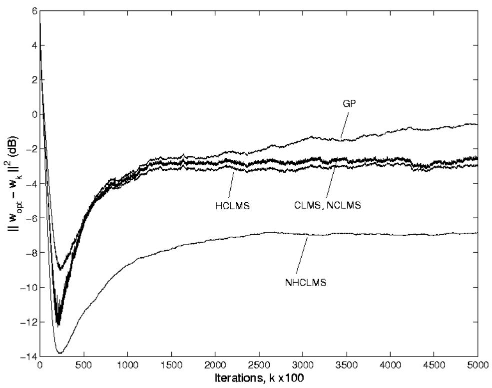

Figure 1: Coefficient-error norm for CLMS, NCLMS, HCLMS, NHCLMS, GP algorithms in 12-bit fixed-point arithmetic.

5. SIMULATION RESULTS

In this section we consider an example where the input signal consists of three sinusoids in white noise. The algorithms are tested in a finite precision environment. The example is taken from [2] and has the input signal given by

$$x_k = \sin(0.3k\pi) + \sin(0.325k\pi) + \sin(0.7k\pi) + n_k \tag{23}$$

where n_k is white Gaussian noise such that the signal-to-noise ratio (SNR) is 40 dB. The filter is constrained to pass the frequency components at 0.1 rad/s and 0.25 rad/s with unity response. This results in four constraints with constraint matrix

$$\mathbf{C}^T = \begin{bmatrix} 1 & \cos(0.2\pi) & \cdots & \cos[(M-1)0.2\pi] \\ 1 & \cos(0.5\pi) & \cdots & \cos[(M-1)0.5\pi] \\ 1 & \sin(0.2\pi) & \cdots & \sin[(M-1)0.2\pi] \\ 1 & \sin(0.5\pi) & \cdots & \sin[(M-1)0.5\pi] \end{bmatrix} \tag{24}$$

and gain vector

$$\mathbf{f}^T = [1 \quad 1 \quad 0 \quad 0] \tag{25}$$

Figure 1 shows the squared deviation of the filter coefficients from their optimal values for the different algorithms. We also show the performance of the gradient-projection (GP) algorithm [1], which is the CLMS algorithm without the correction factor $\mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}(\mathbf{f} - \mathbf{C}^T\mathbf{w})$. The simulation was performed in 12-bit fixed-point arithmetic. The step size used with CLMS, and HCLMS algorithms was equal to 0.1 and with NHCLMS and NCLMS algorithms it was equal to 1. The curves were obtained by averaging over 10 independent trials

As can be seen from the figure, we can clearly see how the NHCLMS algorithm outperforms the other algorithms in terms of steady-state value, whereas the HCLMS performs slightly better than the CLMS and NCLMS algorithms. We can also see that the coefficient-error norm for the GP algorithm grows unbounded as expected [5].

6. CONCLUSIONS

In this paper, we considered orthogonal rotations in the CLMS algorithm and derived a new constrained LMS algorithm and its normalized version. Householder transformations were employed in order to obtain an efficient implementation of the coefficient-vector rotations. In the new algorithm the coefficients are updated only in the subspace defined by the constraints. Therefore, even in finite-precision arithmetic the constraints are satisfied exactly at every iteration. The proposed algorithm also presents lower computational complexity when compared to the conventional CLMS algorithm due to its reduced-dimension update equation. Simulations in finite-precision arithmetic showed superior performance of the new algorithm in terms of coefficient-error norm.

ACKNOWLEDGEMENTS

This work is part of the research project of the Institute of Radio Communication (IRC) funded by Technology Development Center (TEKES), NOKIA Research Center, Sonera, and the Helsinki Telephone Company, Finland, and the research project funded by the Centre-of-Excellence Program (PRONEX—CNPq), Brazil.

7. 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] J. A. Apolinário Jr., S. Werner, and P. S. R. Diniz, "Constrained normalized adaptive filters for CDMA mobile communications," in *Proc. European Signal Processing Conference*, vol. IV, Rhodes, Greece, pp. 2053–2056, 1998.

[4] M. L. R. de Campos, S. Werner, J. A. Apolinário Jr., and T. I. Laakso, "Constrained quasi-Newton algorithm for CDMA mobile communications," in *Proc. International Telecommunications Symposium*, São Paulo, Brazil, pp. 371–376, Aug. 1998.

[5] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas and Propagation*, vol. AP-30, pp. 27–34, Jan. 1982.

[6] G. H. Golub and C. F. Van Loan, *Matrix Computations.* Baltimore: The Johns Hopkins University Press, 1983.