EXTENDED FAST FIXED ORDER RLS ADAPTIVE FILTERS

Ricardo Merched and *Ali H. Sayed*

Department of Electrical Engineering
University of California, Los Angeles, CA 90095

ABSTRACT

Conventional derivations of fast fixed-order RLS filters rely on the shift structure that is characteristic of regressors in a tapped-delay line implementation. In this paper, we study adaptive Laguerre networks, where shift structure no longer holds. We show that fast fixed-order updates are still possible.

1. INTRODUCTION

Fast RLS adaptive schemes are efficient algorithms for updating the least squares solution for growing-length data. While conventional RLS requires $\mathcal{O}(M^2)$ computations per sample, where M is the filter order, its fast versions require only $\mathcal{O}(M)$ operations. Examples of such fast schemes include fast fixed-order filters (e.g., [1, 2, 3]) and also order recursive filters.

The low complexity that is achieved by these algorithms is a consequence of the shift structure that is characteristic of regression vectors in FIR adaptive implementations. Recently, the authors showed that the input data structure that arises from more general networks, such as Laguerre filters, can be exploited to derive fast order-recursive [4] and fast fixed-order implementations [6]. In this paper, we show that fast fixed-order adaptive algorithms can also be derived in explicit form for Laguerre networks. One of the benefits of working with a Laguerre basis is that fewer parameters can be used to model long impulse responses in a stable manner (see e.g., [7, 8]).

2. THE EXTENDED FAST TRANSVERSAL FILTER

Given a column vector $y_N \in \mathbb{C}^{N+1}$ and a data matrix $H_N \in \mathbb{C}^{(N+1)\times M}$, the exponentially-weighted least squares problem seeks the column vector $w \in \mathbb{C}^M$ that solves

$$\min_w \left[\lambda^{N+1}\|\Pi^{-1/2}w\|^2 + \|W_N^{1/2}(y_N - H_N w)\|^2\right] \qquad (1)$$

The matrix Π is a positive-definite regularization matrix, and $W = (\lambda^N \oplus \lambda^{N-1} \oplus \cdots \oplus 1)$. The symbol $*$ denotes complex conjugate transposition. The individual entries of y_N will be denoted by $\{d(i)\}$, and the individual rows of the matrix H_N will be denoted by $\{u_i\}$. The RLS algorithm computes the optimal solution of problem (1) recursively as follows:

$$\begin{aligned}
w_{N+1} &= w_N + g_{N+1}[d(N+1) - u_{N+1}w_N] && (2)\\
g_{N+1} &= \lambda^{-1}P_N u_{N+1}^*\gamma(N+1) && (3)\\
\gamma^{-1}(N+1) &= 1 + \lambda^{-1}u_{N+1}P_N u_{N+1}^* && (4)\\
P_{N+1} &= \lambda^{-1}P_N - g_{N+1}\gamma^{-1}(N+1)g_{N+1}^* && (5)
\end{aligned}$$

with $w_{-1} = 0$ and $P_{-1} = \Pi$. When the regression vectors possess shift structure, it is well known that these recursions can be replaced by more efficient ones. Now, consider the Laguerre network of Fig. 1 where

$$L_0(z) = \frac{\sqrt{1-|a|^2}}{1-az^{-1}} \quad \text{and} \quad L(z) = \frac{z^{-1}-a^*}{1-az^{-1}}, \quad |a| < 1. \qquad (6)$$

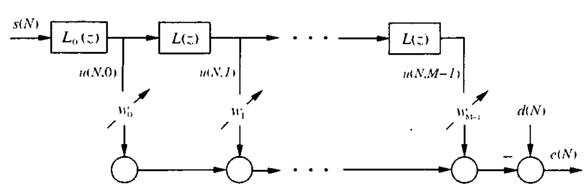

Figure 1: *A transversal Laguerre structure for adaptive filtering.*

The input to the Laguerre filter at time N is denoted by $s(N)$, and the coefficients that combine the outputs of the successive sections $\{L_0(z), L(z)\}$ are denoted by $\{w_k\}$. Using (6) we can relate two successive regression vectors u_N and u_{N+1} as

$$\begin{aligned}
u_{M+1,N} &= [u(N+1,0) \quad u_N] = [u_{N+1} \quad u(N,M-1)]\Phi \\
&= \bar{u}_{M+1,N+1}\Phi && (7)
\end{aligned}$$

where Φ is the $(M+1)\times(M+1)$ matrix

$$\Phi = \begin{bmatrix}
1 & a^* & 0 & 0 & 0 & 0\\
0 & (1-|a|^2) & a^* & 0 & 0 & 0\\
0 & -a(1-|a|^2) & (1-|a|^2) & a^* & 0 & 0\\
0 & a^2(1-|a|^2) & -a(1-|a|^2) & (1-|a|^2) & a^* & 0\\
0 & -a^3 & a^2 & -a & 1 & 0\\
0 & a^4 & -a^3 & a^2 & -a & 1
\end{bmatrix} \qquad (8)$$

This material was based on work supported in part by the National Science Foundation under awards CCR-9732376 and ECS-9820765, and UC Core project CR 98-19. The work of R. Merched was also supported by a fellowship from CAPES, Brazil.

0-7803-6685-9/01/$10.00©2001 IEEE

Note that the regression vectors are not shifted versions of each other anymore. Still, we shall show that a fast RLS algorithm is possible.

Before proceeding, we should remark that since in the paper we also deal with some order-recursive relations, it becomes important to indicate the order of all quantities involved (in addition to a time index). For example, we shall write $w_{M,N}$ instead of w_N to indicate that it is a vector of order M that is computed by using data up to time N.

2.1. Forward Estimation Problem

Consider the input data matrix $H_{M,N}$ and define the coefficient matrix

$$P_{M,N}^{-1} = (\lambda^{N+1}\Pi_M^{-1} + H_{M,N}^* W_N H_{M,N}).$$

Now suppose that one more column is appended to $H_{M,N}$ from the left, i.e.,

$$H_{M+1,N} = \begin{bmatrix} x_{0,N} & H_{M,N} \end{bmatrix} \tag{9}$$

and let

$$P_{M+1,N}^{-1} = (\lambda^{N+1}\Pi_{M+1}^{-1} + H_{M+1,N}^* W_N H_{M+1,N})$$

where $\Pi_{M+1}^{-1} = (\mu \oplus \Pi_M^{-1})$. Then it is straightforward to verify that

$$P_{M+1,N}^{-1} = \begin{bmatrix} \mu\lambda^{N+1} + x_{0,N}^* W_N x_{0,N} & x_{0,N}^* W_N H_{M,N} \\ H_{M,N}^* W_N x_{0,N} & P_{M,N}^{-1} \end{bmatrix}$$

Inverting both sides we get

$$P_{M+1,N} = \begin{bmatrix} 0 & 0 \\ 0 & P_{M,N} \end{bmatrix} + \frac{1}{\zeta_M^f(N)} \begin{bmatrix} 1 \\ -w_{M,N}^f \end{bmatrix} \begin{bmatrix} 1 & -w_{M,N}^{f*} \end{bmatrix} \tag{10}$$

where $w_{M,N}^f$ is the solution to the least-squares problem:

$$\min_{w_M^f} \left[\mu\lambda^{N+1} \|\Pi_M^{-1/2} w_M^f\|^2 + \|W_N^{1/2}(x_{0,N} - H_{M,N} w_M^f)\|^2 \right]$$

whose minimum cost we denote by $\xi_M^f(N)$. This problem projects $x_{0,N}$ onto $\mathcal{R}(H_{M,N})$. Let

$$f_{M,N} = x_{0,N} - H_{M,N} w_{M,N}^f$$

denote the resulting (forward) estimation error vector. Then, the optimal solution for the forward prediction problem admits an RLS recursion of the form

$$w_{M,N}^f = w_{M,N-1}^f + k_{M,N} f_M(N) \tag{11}$$

where $k_{M,N} = g_{M,N}\gamma_M^{-1}(N)$ is the normalized gain vector for a problem of order M, and $f_M(N)$ is the last entry of $f_{M,N}$. We also define the quantity $\zeta_M^f(N) = \mu\lambda^{N+1} + \xi_M^f(N)$.

Substituting Eq. (10) into (3), it is immediate to see that we obtain an order update for $k_{M,N}$ as well,

$$k_{M+1,N} = \begin{bmatrix} 0 \\ k_{M,N} \end{bmatrix} + \frac{\alpha_M^*(N)}{\lambda\zeta_M^f(N-1)} \begin{bmatrix} 1 \\ -w_{M,N-1}^f \end{bmatrix} \tag{12}$$

where $\alpha_M(N)$ is the *a priori* forward prediction error, defined via $f_M(N) = \alpha_M(N)\gamma_M(N)$. A similar order-update can be obtained for the $\gamma_M(N)$, by substituting (10) into (4):

$$\gamma_{M+1}(N) = \gamma_M(N) - \frac{|f_M(N)|^2}{\zeta_M^f(N)}. \tag{13}$$

Combining this recursion with the following time update for the $\zeta_M^f(N)$ (see [4]):

$$\zeta_M^f(N) = \lambda\zeta_M^f(N-1) + \alpha_M^*(N) f_M(N) \tag{14}$$

we obtain an alternative update for $\gamma_M(N)$:

$$\gamma_{M+1}(N) = \gamma_M(N) \frac{\lambda\zeta_M^f(N-1)}{\zeta_M^f(N)}.$$

2.2. Backward Estimation Problem

Similarly to the forward estimation problem, assume that one more column is appended to $H_{M,N}$ from the right, i.e.,

$$\bar{H}_{M+1,N} = \begin{bmatrix} H_{M,N} & x_{M,N} \end{bmatrix} \tag{15}$$

and define the correspondent coefficient matrix as

$$\bar{P}_{M+1,N}^{-1} = (\lambda^{N+1}\bar{\Pi}_{M+1}^{-1} + \bar{H}_{M+1,N}^* W_N \bar{H}_{M+1,N})$$

where

$$\bar{\Pi}_{M+1}^{-1} = \begin{bmatrix} \Pi_M^{-1} & c \\ c^* & \delta \end{bmatrix} \tag{16}$$

for some constant vector c and scalar δ to be specified. Inverting both sides, we obtain:

$$\bar{P}_{M+1,N} = \begin{bmatrix} P_{M,N} & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\zeta_M^b(N)} \begin{bmatrix} -q_N \\ 1 \end{bmatrix} \begin{bmatrix} -q_N^* & 1 \end{bmatrix} \tag{17}$$

This equation has two main differences with respect to the definition of the variables $w_{M,N}^f$ and $\zeta_M^f(N)$, for the forward prediction problem. First, the vector q_N is the sum of two quantities,

$$q_N = w_{M,N}^b + t_N \tag{18}$$

where t_N is given by

$$t_N = \lambda^{N+1} P_{M,N} c \tag{19}$$

The first term of (18) is the solution to the least-squares problem:

$$\min_{w_M^b} \left[\lambda^{N+1} \|\Pi_M^{-1/2} w_M^b\|^2 + \|W_N^{1/2}(x_{M,N} - H_{M,N} w_M^b)\|^2\right]$$

where $\xi_M^b(N)$ is the corresponding minimum cost. This problem projects $x_{M,N}$ onto $\mathcal{R}(H_{M,N})$. The resulting (backward) estimation error vector is given by

$$b_{M,N} = x_{M,N} - H_{M,N} w_{M,N}^b$$

Substituting Eq. (5) into (19), we obtain a recursive relation for t_N (which is analogous to the time-update for $w_{M,N}^b$), and it implies the following time-update for q_N:

$$q_N = q_{N-1} + \eta_M(N) k_{M,N}$$

where $\eta_M(N) = \varepsilon_M(N)\gamma_M(N)$, and

$$\varepsilon_M(N) = \beta_M(N) - u_{M,N} t_{N-1}$$

In addition, the quantity $\zeta_M^b(N)$ is defined by

$$\zeta_M^b(N) \triangleq \xi_M^b(N) + \lambda^{N+1}(\delta - c^* t_N - c^* w_{M,N}^b - w_{M,N}^{b*} c) .$$

Although the update of these terms may look complicated, using the time-update for $w_{M,N}^b$ and t_N, we obtain after some manipulations

$$\zeta_M^b(N) = \lambda \zeta_M^b(N-1) + \varepsilon_M^*(N)\eta_M(N) . \qquad (20)$$

Also, multiplying (17) from the right by $\bar{u}_{M+1,N+1}^*$, we obtain, similar to the forward estimation problem,

$$\begin{bmatrix} k_{M,N} \\ 0 \end{bmatrix} = \bar{k}_{M+1,N} - \nu_M(N) \begin{bmatrix} -w_{M,N-1}^b \\ 1 \end{bmatrix} \qquad (21)$$

where $\nu_M(N) = \varepsilon_M(N)/\lambda\zeta_M^b(N-1)$. The quantity $\nu_M(N)$ is referred to as the *rescue variable* and can be directly obtained as the last entry of $\bar{k}_{M+1,N}$ (to be computed further ahead).

Proceeding similarly to the derivation of (13), we also obtain

$$\bar{\gamma}_{M+1}(N) = \gamma_M(N) - \frac{|\eta_M(N)|^2}{\zeta_M^b(N)} . \qquad (22)$$

Combining Eqs. (20) and (22), it can be shown that

$$\gamma_M(N) = \frac{\bar{\gamma}_{M+1}(N)}{1 - \bar{\gamma}_{M+1}(N)\varepsilon_M(N)\nu_M(N)} .$$

Note that the variables $\varepsilon_M(N)$ and $\eta_M(N)$ play roles similar to the a priori and a posteriori backward prediction problems. Moreover, although all the quantities related to the backward prediction problems satisfy identical recursive equations, here they have different interpretations.

2.3. Exploiting Data Structure

We still need to evaluate $\bar{k}_{M,N}$. For this purpose, we need to identify the variable that is affected by the input data structure. Thus, consider any invertible matrix Φ such as in (8). From Eq. (7), it follows that

$$\bar{H}_{M+1,N+1} = \begin{bmatrix} 0 \\ H_{M+1,N} \end{bmatrix} \Phi^{-1}$$

where $H_{M+1,N}$ and $\bar{H}_{M+1,N+1}$ are the corresponding augmented input data matrices. We then get

$$\begin{aligned} \bar{P}_{M+1,N+1} &= (\lambda^{N+2}\bar{\Pi}_{M+1}^{-1} + \bar{H}_{M+1,N+1}^* W_{N+1} \bar{H}_{M+1,N+1})^{-1} \\ &= (\lambda^{N+2}\bar{\Pi}_{M+1}^{-1} + \Phi^{-*} H_{M+1,N}^* W_N H_{M+1,N} \Phi^{-1})^{-1} \end{aligned}$$

Note that if we could choose

$$\bar{\Pi}_{M+1}^{-1} = \lambda^{-1}\Phi^{-*}\Pi_{M+1}^{-1}\Phi^{-1} \qquad (23)$$

we obtain a simpler relation between $\{\bar{P}_{M+1,N+1}, P_{M+1,N}\}$:

$$\boxed{\bar{P}_{M+1,N+1} = \Phi P_{M+1,N} \Phi^*} \qquad (24)$$

In order for this relation to hold, we need to show how to choose Π_M, c, and δ in order to satisfy (23). Substituting (16) into (23), we get

$$\begin{bmatrix} \Pi_M^{-1} & c \\ c^* & \delta \end{bmatrix} = \lambda^{-1}\Phi^{-*} \begin{bmatrix} \mu & 0 \\ 0 & \Pi_M^{-1} \end{bmatrix} \Phi^{-1} . \qquad (25)$$

Now, the matrix Φ^{-*} can be defined blockwise as

$$\Phi^{-*} = \begin{bmatrix} \bar{v} & \bar{T} \\ 0 & m \end{bmatrix}$$

where

$$m = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$\bar{v} = \begin{bmatrix} a^4 & -a^3 & a^2 & -a & 1 \end{bmatrix}^T$$

Expanding (25), we find that

$$\lambda \Pi_M^{-1} - \bar{T}\Pi_M^{-1}\bar{T}^* = \mu \bar{v}\bar{v}^* . \qquad (26)$$

Hence, if $|a| < \sqrt{\lambda}$, this Lyapunov equation admits a unique positive definite solution Π_M. This is because all the eigenvalues of $\bar{T}$ are either a^* or 0, and the pair $(\lambda^{-1/2}\bar{T}, \bar{v})$ is controllable. From (25), we then obtain

$$\begin{aligned} c &= \lambda^{-1}\bar{T}\Pi_M^{-1} m^* \\ \delta &= \lambda^{-1} m \Pi_M^{-1} m^* = \lambda^{-1}[\Pi^{-1}]_{M-1,M-1} \end{aligned} \qquad (27)$$

From (24), we can now obtain similar relations between $\{g_{M+1,N}, \bar{g}_{M+1,N+1}\}$ and $\{\gamma_{M+1}(N), \bar{\gamma}_{M+1}(N+1)\}$. Thus, multiplying both sides of (24) by $\bar{u}_{N+1}^*$ from the right, we get

$$\bar{g}_{M+1,N+1} = \Phi g_{M+1,N} . \qquad (28)$$

If we further multiply (28) by $\bar{u}_{N+1}$ from the left and subtract 1 from both sides, we have that

$$\bar{\gamma}_{M+1}(N+1) = \gamma_{M+1}(N)$$

which implies the following relation for the normalized extended gain vectors:

$$\boxed{\bar{k}_{M+1,N+1} = \Phi k_{M+1,N}} \tag{29}$$

Initialization

μ is a small positive number; Π is the solution to (26);
c is given by (27).
$\zeta_M^f(-1) = \mu/\lambda$
$\zeta_M^b(0) = \lambda^{-1}[\Pi^{-1}]_{M-1,M-1} - c^*\Pi c$
$w_{M,0} = w_{M,-2}^f = 0$
$q_{-1} = \Pi^{-1}c$

For $N \geq 0$, repeat:

$u(N) = au(N-1) + \sqrt{1-a^2}s(N)$

$\alpha_M(N-1) = u(N) - u_{M,N-1}w_{M,N-2}^f$
$f_M(N-1) = \gamma_M(N-1)\alpha_M(N-1)$
$k_{M+1,N-1} = \begin{bmatrix} 0 \\ k_{M,N-1} \end{bmatrix} + \frac{\alpha_M^*(N-1)}{\lambda\zeta_M^f(N-2)} \begin{bmatrix} 1 \\ -w_{M,N-2}^f \end{bmatrix}$
$\zeta_M^f(N-1) = \lambda\zeta_M^f(N-2) + \alpha_M^*(N-1)f_M(N-1)$
$w_{M,N-1}^f = w_{M,N-2}^f + k_{M,N-1}f_M(N-1)$
$\bar{\gamma}_{M+1}(N) = \gamma_M(N-1)\frac{\lambda\zeta_M^f(N-2)}{\zeta_M^f(N-1)}$

$\bar{k}_{M+1,N} = \Phi k_{M+1,N-1}$
$\nu_M(N) =$ (last entry of $\bar{k}_{M+1,N}$)
$k_{M,N} = \bar{k}_{1:M,N} + \nu_M(N)q_{N-1}$
$\varepsilon_M(N) = \lambda\zeta_M^b(N-1)\nu_M^*(N)$
$\gamma_M(N) = \frac{\bar{\gamma}_{M+1}(N)}{1-\bar{\gamma}_{M+1}(N)\varepsilon_M(N)\nu_M(N)}$
$\eta_M(N) = \gamma_M(N)\varepsilon_M(N)$
$\zeta_M^b(N) = \lambda\zeta_M^b(N-1) + \varepsilon_M^*(N)\eta_M(N)$
$q_N = q_{N-1} + k_{M,N}\eta_M(N)$

$\epsilon_M(N) = d(N) - u_{M,N}w_{M,N-1}$
$e_M(N) = \gamma_M(N)\epsilon_M(N)$
$w_{M,N} = w_{M,N-1} + k_{M,N}e_M(N)$

Table 1: The extended fast fixed-order RLS filter for Laguerre networks.

This relation shows that the time update of the gain vector $k_{M,N}$, which is necessary to update the optimal solution

$$w_{M,N+1} = w_{M,N} + k_{M,N+1}e_M(N+1)$$

can be efficiently performed in three main steps:

1) Order update $k_{M,N} \rightarrow k_{M+1,N}$ [Eq. (12)];
2) Time update $k_{M+1,N} \rightarrow \bar{k}_{M+1,N+1}$ [Eq. (29)];
3) Order downdate $\bar{k}_{M+1,N+1} \rightarrow k_{M,N+1}$ [Eq. (21)].

Note that when $a = 0$, we have $\Phi = I$ and therefore $\bar{k}_{M+1,N+1} = k_{M+1,N}$, in which case the recursions collapse to the FTF algorithm [2]. Equation (29) is the only recursion that uses the fact that the input data has structure. For example, for Laguerre-based filters, this multiplication is essentially a convolution, which can be performed with $\mathcal{O}(M)$ operations. The cost of the usual FIR FTF algorithm is known to be $\mathcal{O}(7M)$ operations. The overall cost for the Laguerre case simply amounts to $\mathcal{O}(8M)$ operations. Table 1 shows the resulting generalized FTF algorithm.

We should mention that the general algorithm proposed here can face some stability problems just like the standard FTF algorithm [2]. A stabilization procedure can be pursued along the lines of [9], and this will be developed elsewhere. In addition, array algorithms are developed in [5, 6].

3. REFERENCES

[1] G. Carayannis, D. Manolakis, and N. Kalouptsidis, "A fast sequential algorithm for least squares filtering and prediction," *IEEE Trans. on Acoust., Speech, Signal Proc.*, vol. ASSP-31, pp. 1394–1402, December 1983.

[2] J. Cioffi and T. Kailath, "Fast recursive-least-squares transversal filters for adaptive filtering," *IEEE Trans. on Acoust., Speech Signal Processing*, vol. ASSP-32, pp. 304-337, April 1984.

[3] A. H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Magazine*, vol. 11, no. 3, pp. 18-60, Jul. 1994.

[4] R. Merched and A. H. Sayed, "Order-recursive RLS Laguerre adaptive filtering," *IEEE Transactions on Signal Processing*, vol. 48, no. 11, pp. 3000-3010, Nov. 2000.

[5] R. Merched and A. H. Sayed, "Orthonormal realization of fast fixed-order RLS adaptive filters," *Proc. ICASSP*, Salt Lake City, Utah, May 2001.

[6] R. Merched and A. H. Sayed, "Fast RLS Laguerre adaptive filtering," *Proc. Allerton Conference*, IL, Sep. 1999.

[7] B. Wahlberg, "System identification using Laguerre models," *IEEE Trans, Automat. Control.*, vol. 36, pp. 551-562, May. 1991.

[8] J. W. Davidson and D. D. Falconer, "Reduced complexity echo cancelation using orthonormal functions," *IEEE Trans. on Circuits Syst.*, vol. 38, no. 1, pp. 20-28, Jan. 1991.

[9] D. T. M. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least-squares adaptive filtering," *IEEE Trans. on Signal Processing*, vol. 39, pp. 92-113, Jan. 1991.