Laguerre-domain adaptive filters

den Brinker, A.C.

Published in:
IEEE Transactions on Signal Processing

DOI:
10.1109/78.285660

Published: 01/01/1994

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
to be essentially independent of the influence of phase error. A similar effect may be seen in the time domain results presented in [6] for a sampling rate four times the input frequency ($\gamma = 90^\circ$). Figs. 2 and 3 also depict the maximum stable value of convergence coefficient for time delays of one sample and one cycle. Observe, especially in Fig. 3, that the characteristics of the relationship between the phase error and maximum stable convergence coefficient have changed. In addition, with the time delays, the peaks near the $\pm 90^\circ$ bound are removed, and so the shape of the curve approximates the cosine function shape of the complex algorithm, as suggested in [6]. However, the curve is still not, in general, symmetric about the $\phi = 0^\circ$ point, the exception being at a sampling rate four times the input frequency ($\gamma = 90^\circ$) used in [6]. Although these convergence characteristics of the time domain-filtered $z$ LMS algorithm with a cancellation loop transfer function phase estimation error are somewhat interesting, they are also somewhat discouraging from the viewpoint of the stated objective of the analysis. In fact, it is almost impossible to provide a more quantitative assessment of the effect of transfer function phase estimation error beyond stating that the tolerable bounds of this error are $|\phi| < 90^\circ$.

IV. CONCLUSIONS

Errors in the estimation of the cancellation path transfer function for adaptive noise and vibration control systems implementing the filtered-$z$ LMS algorithm will have an influence on the stability of the algorithm. Errors in the estimation of the magnitude of this quantity will alter the maximum stable value of convergence coefficient through an inverse proportional relationship. It can further be said that it is possible for the algorithm to be made stable, provided the error in the estimation of the phase of the transfer function does not exceed $\pm 90^\circ$. The effect that a phase error has on algorithm stability between these bounds is more difficult to predict, owing to the alteration of the error surface as seen by the algorithm induced by the phase errors. The effect will normally not be symmetric about the $0^\circ$ phase error point and may, in fact, cause the stability of the algorithm to increase for some values of error. This is in contrast to the case where the complex algorithm is used, where the maximum stable value of convergence coefficient is simply reduced by a factor proportional to the cosine of the phase error.

REFERENCES


Laguerre-Domain Adaptive Filters

Albertus C. den Brinker

Abstract—Using a tapped-delay-line as an adaptive filter, the complexity of the filter increases with increasing correlation length of input and reference signal. We seek simple adaptive filter structures such that for a long correlation length only a minor complexity of the filter is needed. We consider adaptive mechanisms governed by an exponentially weighted squared-error criterion. Laguerre-domain adaptive filters are introduced, which leads to a tapped IIR-filter line. These filters contain a discount factor as a free variable, which makes it possible to set the memory and the number of adaptive coefficients independently. Convergence properties of the proposed adaptive filters are discussed.

I. INTRODUCTION

The following problem is considered. Suppose we have a reference signal $r(n)$ where part of $r(n)$, say $r_1(n)$, is correlated with another signal $x(n)$. The correlation is considered to be a consequence of some linear causal filtering. This filtering is described by an impulse response called $f(n)$ and therefore, $r_1(n) = f(n) * x(n)$, where $*$ denotes convolution. We want to construct a filter that approximately performs the filter operation $f(n)$. We allow linear, causal, slowly time-varying impulse responses $f(n)$.

Usually, a transversal filter is used in these cases. However, in the case of a long correlation length, i.e., $f(n)$ has a long nonnegligible tail, we need a transversal filter with many taps. Thus, the complexity of the adaptive filter, defined as the number of adaptive parameters, increases with increasing correlation length for a proper functioning of the adaptive filter. On the other hand, a large number of parameters is of course cumbersome to control and, in itself, deteriorates the performance.

Starting from a more general adaptive filter and an exponentially weighted squared-error criterion, we seek appropriate filter structures that are able to cope with the situation of a long correlation length without the need to increase complexity. It is shown that what we call Laguerre-domain adaptive filters (LDAF) are suited for this job. The filter coefficients can be updated using a recursive least-squares algorithm.

II. THE OPTIMIZATION CRITERION

Consider the adaptive filter $F$ shown in Fig. 1. We have an input signal $x(n)$ and a desired response or reference signal $r(n)$. The adaptive filter $F$ has an output signal called $y(n)$, and we construct the error signal $e(n)$ according to $e(n) = r(n) - y(n)$.

The filter parameters of $F$ are called the weights $w_m$. Our aim is to set these weights in such a way that the error is minimal in a certain sense. We consider minimization of the exponentially weighted squared error $J(n)$:

$$J(n) = \sum_{k=-\infty}^{n} \{f(k)\}^2 e^{-\theta k}, \quad 0 < \theta < 1.$$  \hspace{1cm} (1)

Manuscript received April 2, 1992; revised April 22, 1993. The associate editor coordinating the review of this paper and approving it for publication was Dr. Hong Fan.

The author is with the Electrical Engineering Department, Eindhoven University of Technology, Eindhoven, The Netherlands.

IEEE Log Number 9215271.
We assume a linear regression model where the filter output \( y(n) \) is written as
\[
y(n) = \sum_{m=0}^{M} w_m u_m(n),
\] (2)
The signals \( u_m(n) \) are internal signals of the adaptive filter \( F \), and there are \( M+1 \) weights indexed 0 to \( M \). The signals \( u_m(n) \) are derived from \( x \) by some linear causal filter operation. In essence the adaptive filter \( F \) is a filter bank, not necessarily the usual tapped-delay-line.

Taking the derivatives of this optimization criterion \( J(n) \) with respect to the filter weights \( u_m \), setting these to zero and using (2) gives
\[
\sum_{m=0}^{M} \sum_{k=-\infty}^{n} w_k (k) u_m(k) g_m^0 - k = \sum_{k=-\infty}^{n} r(k) u_m(k) g_m^0 - k.
\] (3)
This equation holds for all \( m \) in the optimal situation. In (3), we derived from \( J(n) \) the input signal \( z(n) \) to be filtered. The derived normal equation is equal to a deterministic normal equation for the coefficients \( u_m \) of the adaptive filter. The relation \( H H^T w = H \Phi \), where \( w = [w_0, w_1, \ldots, w_M]^T \), the Laguerre spectrum of \( r \), and \( H \) is a matrix containing rowwise the Laguerre spectra of the windowed internal filter signals \( u_m \) (0 \( \leq m \leq M \))
\[
H = \begin{bmatrix}
w_0 & 0 & 0 & \cdots \\
0 & w_1 & 0 & \cdots \\
0 & 0 & w_2 & \cdots \\
\vdots & \vdots & \vdots & \ddots \\
0 & 0 & 0 & \cdots \\
1 & 1 & 1 & \cdots \\
\cdots & \cdots & \cdots & \ddots \\
M & M & M & \cdots 
\end{bmatrix}.
\] (9)
The relation \( H H^T y = H \phi \) is the normal equation for the coefficients of the adaptive filter. Derived normal equation is equal to a deterministic normal equation for a tapped delay line (cf. [3]). The matrix \( H H^T \) contains the coefficients \( c_0 \) (see [1] and [2]) of the local cross-correlation functions of \( u_0 \) and \( u_m \), whereas the vector \( H \phi \) contains the coefficients \( c_0 \) of the local cross-correlation functions of \( u_m \) and \( r \)
equation for the bias proposed adaptive filter. For this situation three statistical properties process. A reference signal linearly combined in an output signal filter bank is combined with variable weights and constitutes the \( \tilde{\psi}(n) \) according to (10), with \( \psi(n) \) states that we do not need to perform a prefiltering operation on \( \tilde{\psi}(n) \).\( \psi(n) \) hold for finite \( M \) matrices \( H(n)H^T(n) \) together with vector \( u \) of length \( M \). Note that (15) states that we do not need to perform a decomposition of the reference signal \( r \) in order to establish the vector \( H \). Using the matrix inversion lemma (e.g., [4]), this leads to the well-known RLS algorithm (P(n) being the inverse of \( H(n)H^T(n) \))

\[
\tilde{\psi}(n) = \frac{P(n-1)\tilde{\psi}(n)}{\theta + u^T(n)P(n-1)\tilde{\psi}(n)}
\]

\[
\alpha(n) = r(n) - \tilde{\psi}(n)\tilde{\psi}(n)^T(n),
\]

\[
\psi(n) = \psi(n-1) + \psi(n)\alpha(n),
\]

\[
P(n) = \theta^{-1}P(n-1) - \theta^{-1}\tilde{\psi}(n)\tilde{\psi}(n)^T(n)P(n-1).
\]

\[
(16)
\]

\[
(17)
\]

\[
(18)
\]

\[
(19)
\]

IV. UPDATING MECHANISM

The exponentially windowed error criterion (1) automatically gives rise to the following two recursive relations

\[
H(n)H^T(n) = \theta H(n-1)H^T(n-1) + \psi(n)\psi(n)^T(n),
\]

\[
H(n) = \theta H(n-1)g(n-1) + \psi(n)r(n)
\]

which hold for finite \( M \times M \) matrices \( H \) together with vector \( u \) of length \( M \). Note that (15) states that we do not need to perform a decomposition of the reference signal \( r \) in order to establish the vector \( H \). Using the matrix inversion lemma (e.g., [4]), this leads to the well-known RLS algorithm (P(n) being the inverse of \( H(n)H^T(n) \))

\[
\tilde{\psi}(n) = \frac{P(n-1)\tilde{\psi}(n)}{\theta + u^T(n)P(n-1)\tilde{\psi}(n)}
\]

\[
\alpha(n) = r(n) - \tilde{\psi}(n)\tilde{\psi}(n)^T(n),
\]

\[
\psi(n) = \psi(n-1) + \tilde{\psi}(n)\alpha(n),
\]

\[
P(n) = \theta^{-1}P(n-1) - \theta^{-1}\tilde{\psi}(n)\tilde{\psi}(n)^T(n)P(n-1).
\]

V. CONVERGENCE PROPERTIES OF THE LDAF

Consider the situation in Fig. 2. We have two filter banks \( F \) according to (10), with \( F_0(z) \) unspecified and finite \( M \). Both filter banks have input signal \( x(n) \). The outputs of one of these are linearly combined in an output signal \( r_1(n) \) by constant weights \( \tilde{w}_m \) \( (0 \leq m \leq M) \). The process \( x(n) \rightarrow r_1(n) \) is thus a stationary process. A reference signal \( r(n) \) is produced as a summation of \( r_1(n) \) and a statistically independent noise signal \( \tilde{\psi}(n) \). The other filter bank is combined with variable weights and constitutes the proposed adaptive filter. For this situation three statistical properties are derived.

From (18) and \( \psi(n) = \tilde{\psi}(n) + \tilde{\psi}(n) \), we construct the updating equation for the bias

\[
\tilde{\psi}(n) = \tilde{\psi}(n-1) + \tilde{\psi}(n)\alpha(n).
\]

\[
(20)
\]

Substitution of (17) in (20) and using \( \tilde{\psi}(n) = P(n)\psi(n) \) and (14) gives the following time-variant difference equation for \( \tilde{\psi}(n) \):

\[
\tilde{\psi}(n) - \theta P(n)H(n-1)H^T(n-1)\tilde{\psi}(n-1) = P(n)\psi(n)\tilde{\psi}(n).
\]

(21)

This equation is linear in \( \tilde{\psi}(n) \) provided that \( \tilde{\psi}(n) \) and \( r(n) \) are statistically independent. Since \( \psi(n) \), \( H(n-1)H^T(n-1) \), and \( P(n) \) are proportional, proportional to the square, and inversely proportional to the square of the amplitude of \( x(n) \), respectively, the bias is inversely proportional to the amplitude of the input signal \( x(n) \).

Suppose now that we start at time \( n = 1 \) with \( \tilde{\psi}(0) = \tilde{\psi}_0 \) and matrix \( H(0)H^T(0) \). (The initial matrix \( H(0)H^T(0) \) should be taken nonsingular. In RLS algorithms, one commonly takes \( H(0)H^T(0) = \delta I \), where \( I \) is the identity matrix, and \( \delta \) is some small positive constant [4].) Solving the difference equation (21) gives

\[
\tilde{\psi}(n) = \theta^n P(n)H(0)H^T(0)\tilde{\psi}_0 + P(n)\sum_{k=1}^n \theta^{n-k}\tilde{\psi}(k)r_2(k).
\]

(22)

Suppose now that \( x(n) \) is a deterministic signal and thus that \( \psi(k) \) and \( P(n) \) are deterministic quantities. Suppose furthermore that \( r_2(n) \) is a stochastic signal derived from a zero-mean process. Consequently, the bias \( \tilde{\psi}(n) \) is a stochastic signal and its expectation \( E[\tilde{\psi}(n)] \) is

\[
E[\tilde{\psi}(n)] = \theta^n P(n)H(0)H^T(0)\tilde{\psi}_0
\]

(23)

The foregoing equation expresses that if \( P(n) \) increases slower than exponentially, the expectation \( E[\tilde{\psi}(n)] \) tends to zero for \( n \to \infty \). In fact, (23) expresses the identifiability of the unknown process given a certain input signal \( x(n) \). One can always construct signals \( x(n) \) such that \( \tilde{\psi}(n) \) does not tend to \( \tilde{\psi}(n) \) for \( n \to \infty \).

Given a proper input signal such that the unknown process is identifiable, the weights \( \tilde{\psi}(n) \) will fluctuate around the optimal solution \( \tilde{\psi}(n) \) after a first learning period. The difference between the actual and optimal solution is again denoted as \( \tilde{\psi}(n) \) and is called the weight-error vector. We calculate the weight-error correlation matrix \( E[\tilde{\psi}(n)\tilde{\psi}(n)^T] \).

We assume that the adaptive filter is working in its steady state and that the signals \( x(n) \) and \( r_2(n) \) extend from \( n = -\infty \) to \( +\infty \). We assume that \( x(n) \) is a deterministic signal and that \( r_2(n) \) is a wide-sense stationary, zero-mean process with variance \( \sigma^2 \). Repeated application of (21) gives

\[
E[\tilde{\psi}(n)\tilde{\psi}(n)^T] = \sigma^2 P(n)\sum_{k=-\infty}^{n} \theta^{n-k}\tilde{\psi}(k)\tilde{\psi}(k)^T P(n).
\]

(24)

The covariance matrix is proportional to the variance of \( r_2(n) \). Since \( \tilde{\psi}(n) \) is inversely proportional to the amplitude of the input signal \( x \), the covariance matrix is inversely proportional to the square of this amplitude.

As a next item, we consider the mean-squared innovation \( J'(n) \) defined as \( J'(n) = E[\alpha(n)^2] \). We again consider the steady state situation, and we assume that \( r_2(n) \) is a wide-sense stationary, zero-mean, white-noise process with variance \( \sigma^2 \). Since \( r_2(n) \) is assumed to be a white-noise process, \( r_2(n) \) and \( \tilde{\psi}(n-1) \) are statistically independent. By using \( \alpha(n) = r_2(n) - \tilde{\psi}(n-1)\tilde{\psi}(n) \) and (24) we find

\[
J'(n) = \sigma^2 + \sigma^2 E[\tilde{\psi}(n)\tilde{\psi}(n)^T] P(n)\sum_{k=-\infty}^{n} \theta^{n-k}\tilde{\psi}(k)\tilde{\psi}(k)^T P(n).
\]

(25)

From the foregoing relation, we conclude that the mean-squared innovation is proportional to the variance of \( r_2(n) \) and independent of the amplitude of the input signal.

We have simulated the adaptive filter in the situation shown in Fig. 2. The signals \( x(n) \) and \( r_2(n) \) were derived from two
Laguerre filter are equal (Section IV), the differences between these not actually use the information provided by the local analysis in the described as Laguerre series.

We addressed in the introduction.

between input and reference signal is (slowly) time varying. The Laguerre-domain adaptive filter is expected to perform well for a versions of the input and reference signals. Windowing of the data is required since in adaptive filtering it is assumed that the correlation of fluctuations yields larger fluctuations in the innovation. The conclusion is that the Laguerre-domain adaptive filter is expected to perform well for a small number of weights and a large $\theta$. This is exactly the problem we addressed in the introduction.

VI. DISCUSSION

We developed an adaptive filter starting by considering windowed versions of the input and reference signals. Windowing of the data is required since in adaptive filtering it is assumed that the correlation between input and reference signal is (slowly) time varying. An exponential window was chosen, and consequently, the signals were described as Laguerre series.

Our starting point differs from the usual approach in system identification using Laguerre functions [3]-[12]. There one commonly starts by a description of the model by a Laguerre filter, whereas we started by transforming the windowed signals to the Laguerre domain. This provides a local analysis of the input signal. We did not actually use the information provided by the local analysis in the present study. As a consequence of this and the fact that the model sets of the introduced adaptive filter and the more commonly used Laguerre filter are equal (Section IV), the differences between these two approaches should be sought in aspects of implementation.

There are two degrees of freedom in the Laguerre domain adaptive filter. The first one is the initial filtering stage $F_0$. Taking $F_0(z) = z \sqrt{1 - \theta} / (z - \theta)$ is equal to taking a truncated Laguerre filter. Other choices for $F_0$ can be interpreted as prefiltering the input data $x$ and subsequently using a truncated Laguerre filter. In this way, one can shape the spectrum of the input signal on the basis of a priori information in order to optimally exploit the properties of a truncated LDAF with a minimal number of coefficients, e.g., if it is known that we are approximating a process with a high-frequency falloff, one can already suppress these parts of the input signal by $F_0$.

The second degree of freedom is the parameter $\theta$, which is in essence a scale factor determining how long the memory of the filter is taken. Although each stable linear causal time-invariant operator can be approximated by a Laguerre filter, the discount factor $\theta$ must be carefully chosen in order to obtain a minimal number of meaningful coefficients in the adaptive filter. In the case that the unknown process has a broadband spectrum, one can resort to multiscale adaptive filtering [12].

Current research includes further software simulations of the LDAF. Hardware implementation of the proposed adaptive filter is considered as well.

REFERENCES


Least-Squares Design of Higher Order Nonrecursive Differentiators

S. Sunder and Ravi P. Ramachandran

Abstract—A method is described that can be used to design nonrecursive linear-phase higher order differentiators that can perform differentiation over any frequency range. The method is based on formulating the absolute mean-square error between the amplitude responses of the practical and ideal differentiator as a quadratic function. The coefficients of the differentiators are obtained by solving a set of linear equations. This method leads to a lower mean-square error and is computationally more efficient than both the eigenfilter method and the method based on the Remez exchange algorithm. Design of differentiators based on minimization of the relative mean-square error is also carried out. Finally, our method is extended to the design of frequency selective higher order differentiators.

Manuscript received April 24, 1992; revised March 17, 1993. The associate editor coordinating the review of this paper and approving it for publication was Prof. Tamal Bose.

S. Sunder is with the Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada.
R. P. Ramachandran is with the Caip Center, Department of Electrical Engineering, Rutgers University, Piscataway, NJ 08855.
IEEE Log Number 9215275.

1053-587X/94$4.00 © 1994 IEEE