A Note on the Optimal Expansion of Volterra Models using Laguerre Functions

Automatica (Oxford)

Vol. 42, no. 4, pp. 689-693, 2006

Unformatted Manuscript
A NOTE ON THE OPTIMAL EXPANSION OF VOLterra MODELS USING LAGUERRE FUNCTIONS

Ricardo J. G. B. Campello\textsuperscript{a,b}, Wagner C. do Amaral\textsuperscript{b}, Gérard Favier\textsuperscript{c}

\textsuperscript{a}COPOP/UNISANTOS, R. Dr. Carvalho de Mendonça 144, CEP 11070-906, Santos - SP, Brazil
\textsuperscript{b}DCA/FEEC/UNICAMP, CEP 13083-970, Campinas - SP, Brazil
\textsuperscript{c}Laboratoire I3S/CNRS/UNSA, B.P. 121, 06903 Sophia Antipolis Cédex, France

Abstract

This work tackles the problem of expanding Volterra models using Laguerre functions. A strict global optimal solution is derived when each multidimensional kernel of the model is decomposed into a set of independent orthonormal bases, each of which parameterized by an individual Laguerre pole intended for representing the dominant dynamic of the kernel along a particular dimension. It is proved that the solution derived minimizes the upper bound of the square norm of the error resulting from the practical truncation of the Laguerre series expansion into a finite number of functions. This is an extension of the results in [5], where an optimal solution was obtained for the usual yet particular case in which a single Laguerre pole is used for expanding a given kernel along all its dimensions. It is also proved that the particular and extended solutions are equivalent to each other when the Volterra kernels are symmetric.

Key words: Non-linear Systems; Volterra Series; Laguerre Functions; Optimization; Model Reduction.

1 Introduction

The use of Volterra models for modeling, identification and control of non-linear systems has been a subject of great interest for decades [8,13,7]. An \textit{M}-th-order Volterra model is generically described in discrete time as:

\begin{equation}
    y(k) = \sum_{m=1}^{M} \sum_{k_1=0}^{\infty} \cdots \sum_{k_m=0}^{\infty} h_m(k_1,k_2,\cdots,k_m) \prod_{j=1}^{m} u(k-k_j)
\end{equation}

where \( u, y \) and \( h_m \) are the input, the output and the \( m \)-th-order kernel, respectively. This model is a non-linear generalization of the well-known impulse response model [1] and allows to approximate to desired accuracy a wide class of real-world systems, namely, causal systems with fading memory and bounded input [2]. The central problem in practice is how to select Volterra kernels that provide an adequate representation of the system to be modeled. The main drawback is that the kernels are, in principle, non-parameterized functions whose measurement is possible only if their individual contributions can be separated from the total system response [14].

When modeling stable systems, a straightforward approach to the above-mentioned problem is to ignore the potentially negligible long memory terms of the kernels, i.e. \{ \( h_m(k_1,\cdots,k_m) : \ k_j > \varepsilon_m \) \}, where \( \varepsilon_m \) is a sufficiently big positive integer, in such a way that the remaining terms can be treated as individual parameters to be estimated [7]. Since model (1) is linear in these parameters, classical estimation algorithms like the Recursive Least Squares (RLS) can be applied. This approach, however, usually makes the model over-parameterized.

A usual and effective strategy to reduce the parametric complexity of the model, first suggested by Wiener [18], is the expansion of the Volterra kernels using orthonormal basis functions. The most commonly used orthonormal basis is the Laguerre basis [3,10,17], for the corresponding functions are completely determined by a single real valued parameter: the Laguerre pole. This parameter plays a key role in the design of parsimonious Laguerre models since the number of basis functions needed to approximate the system to desired accuracy directly depends on it.

Although the systematic selection of the Laguerre pole according to some optimization criterion is a problem that has been investigated since the 1960s [6,11,9,12,16,15], the obtained mathematical results essentially concern linear models. Nevertheless, in a previous work by the authors [5] an analytic strict global solution for the optimal Laguerre series

Draft
The transfer functions of the Laguerre filters are given by:

\[ y(k) = \sum_{m=1}^{M} \sum_{l=1}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \alpha_{i,j,m} \prod_{j=1}^{m} \phi_{m,j}(k) \]

where \( \phi_{m,j}(k) \) is the \( j \)-th function of the Laguerre basis responsible for expanding the \( m \)-th-order kernel along its \( j \)-th dimension, and \( \alpha_{i,j,m} \) are the corresponding expansion coefficients.

From equations (1) and (2), the Volterra model can easily be rewritten as:

\[ y(k) = \sum_{m=1}^{M} \sum_{l=1}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \alpha_{i,j,m} \prod_{j=1}^{m} \phi_{m,j}(k) \]

with \( \Phi_{m,j}(\tau) \) denoting the output of a Laguerre filter, i.e., 

\[ \Phi_{m,j}(\tau) = \sum_{k=0}^{\infty} \phi_{m,j}(\tau) \]

The transfer functions of the Laguerre filters are given by:

\[ \Phi_{m,j}(z) = Z\{\phi_{m,j}(\tau)\} = \frac{\sqrt{1 - p_{m,j}^2}}{z - p_{m,j}} \left( 1 - \frac{p_{m,j}^2}{z - p_{m,j}} \right)^{l-1} \]

where \( Z \) stands for the unilateral \( z \)-transform and \( p_{m,j} \in (-1,1) \) is the stable pole which parameterizes the basis related to the \( j \)-th dimension of the \( m \)-th order kernel.

In fact, equations (2) and (3) are exact under the only condition that \( h_m \) is stable, even when a single Laguerre basis is used along all the dimensions of the kernel \( (p_{m,1} = p_{m,2} = \cdots = p_{m,m}) \). In practice, however, the series expansion is truncated into a finite number of functions. In this case, using an independent basis for each dimension enhances flexibility and is expected to reduce the truncation error when the dominant dynamics of the kernel along its multiple dimensions are different from one another.

## 3 Optimization of the Laguerre Poles

The underlying problem considered here is how to select the Laguerre poles \( p_{m,j} \) in (5) so as to minimize the error resulting from the truncation of the series expansion (2) into a finite number of terms. This problem can be tackled indirectly by minimizing the following cost function with respect to the Laguerre poles:

\[ J_m = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (i_1 + \cdots + i_m) \alpha^2_{i_1,\ldots,i_m} \]

This function forces a fast convergence of the orthonormal series by linearly increasing the cost assigned to each additional Laguerre coefficient \( \alpha_{i} \). Thus, minimizing (6) implicitly implies the minimization of the error that results from a truncation of the series into an arbitrary number of functions or, equivalently, the minimization of the number of functions needed to approximate the kernel with arbitrary error.

**Theorem 1** Let the \( m \)-th order Volterra kernel satisfy the stability and unit delay conditions, i.e.

1. \( \sum_{k=0}^{\infty} \sum_{k_m=0}^{\infty} |h_m(k_1,\ldots,k_m)| < \infty \)
2. \( h_m(k_1,\ldots,k_m) = 0 \) if \( \exists l \in \{1,\ldots,m\} : k_l = 0 \).

Then, the cost function in (6) can be rewritten as:

\[ J_m = \sum_{l=1}^{m} \left( (S_{1,l} - 1) p_{m,l}^2 + (S_{2,l} - 2 S_{1,l} + 1) p_{m,l} + S_{1,l} \right) \]

where \( S_{1,l} \) and \( S_{2,l} \) are constant terms which depend exclusively on the \( m \)-th order Volterra kernel, as follows:

\[ |h_m|^2 = \sum_{k_1=0}^{\infty} \cdots \sum_{k_m=0}^{\infty} h_m(k_1,\ldots,k_m) \]

Clearly, equation (6) is not the only way to force a fast convergence of the series, but the mathematical properties of alternative cost functions are still to be investigated, even in the linear case.
\[ S_{1,l} = \frac{1}{||h_m||^2} \sum_{k_1=0}^{\infty} \cdots \sum_{k_m=0}^{\infty} k_1 h_m^2(k_1, \cdots, k_m) \]  
(9)

\[ S_{2,l} = \frac{1}{||h_m||^2} \sum_{k_1=0}^{\infty} \cdots \sum_{k_m=0}^{\infty} k_1 [\Delta h_m(k_1, \cdots, k_m)]^2 \]  
(10)

where \( l = 1, \cdots, m \) and the operator \( \Delta_l \) is defined as \( \Delta h_m(k_1, \cdots, k_m) = h_m(k_1, \cdots, k_l + 1, \cdots, k_m) - h_m(k_1, \cdots, k_l, k_{l+1}, \cdots, k_m) \).

**Proof.** See the Appendix. \( \square \)

By using theorem 1 it is possible to select the poles \( p_{m,l} \) (\( l = 1, \cdots, m \)) which parameterize the Laguerre expansion of the \( m \)th-order kernel \( h_m \) by solving the following problem:

\[
\min_{p_m \in \mathcal{P}} J_m = \frac{J_m}{||h_m||^2} \quad (11)
\]

where \( \mathcal{P} = \{ p_{m,1}, \cdots, p_{m,m} \}^T \), \( \mathcal{P} \triangleq \{ p_{m,l} \in \mathbb{R}^m : |p_{m,l}| < 1 \} \) and \( J_m \) is given by equation (7).

**Theorem 2** If kernel \( h_m \) satisfies the conditions stated in theorem 1, then problem (11) has a strict global optimal solution given by:

\[
p_{m,l}^* = \frac{2S_{1,l} - S_{2,l}}{2S_{1,l} - 1 + \sqrt{4S_{1,l}S_{2,l} - S_{2,l}^2 - 2S_{2,l}}} \quad (12)
\]

for \( l = 1, \cdots, m \).

**Proof.** Note from (7) that the cost function \( J_m \) in (11) is a sum of \( m \) independent terms, each of which depending exclusively on a single Laguerre pole \( p_{m,l} \) (\( l = 1, \cdots, m \)). Hence, since the minimum of the sum of a set of independent terms is the sum of their individual minima, problem (11) can be solved separately for each Laguerre pole, i.e.,

\[
\min_{|p_{m,l}| < 1} \left( \frac{(S_{1,l} - 1)(p_{m,l}^2 + (S_{2,l} - 2S_{1,l} + 1)p_{m,l} + S_{1,l})}{1 - p_{m,l}^2} \right) \quad (13)
\]

This problem consists of the scalar minimization of a pseudo-convex function within an open convex domain and can be shown to have a strict global optimal solution given by equation (12) [4,5]. \( \square \)

**Theorem 3** The optimal Laguerre poles given by equation (12) minimize the upper bound of the squared norm of the error resulting from the truncation of the series expansion into a finite number \( N_m \) of functions.

**Proof.** Truncating the kernel expansion given by equation (2) into \( N_m \) Laguerre functions yields

\[
h_m(k_1, \cdots, k_m) = \sum_{i=1}^{N_m} \cdots \sum_{l=1}^{N_m} a_{i, \cdots, l} \prod_{j=1}^{m} \phi_{m,j,i,j}(k_j).
\]

The proof is straightforward since the terms \( S_{1,l} \) and \( S_{2,l} \) in (9) and (10), respectively, are constant for \( l = 1, \cdots, m \) when \( h_m \) is symmetric. In this case, the single solution in (17) becomes clearly equivalent to every individual solution in (12), i.e., \( p_{m}^* = p_{m,1}^* = p_{m,2}^* = \cdots = p_{m,m}^* \). \( \square \)

Note that, under the perspective of the output of the Volterra model in (1), it is possible to replace any non-symmetric...
kernel with a symmetric equivalent by means of an ordinary symmetrization procedure [14]. However, the symmetric kernels are equivalent to their non-symmetric counterparts only in terms of the model output, i.e., they are not equivalent to each other as multidimensional functions to be described by means of a truncated series expansion. This means that the symmetrization of a given kernel does not ensure that its conventional expansion using a single Laguerre basis is equivalent to the expansion of the original kernel using a particular basis for each dimension.

In summary, theorem 4 suggests that: (i) if one begins by identifying a symmetric Volterra model, as usual in practice, then there is no reason to use multiple Laguerre bases to approximate each kernel; and (ii) if one does decide to adopt multiple bases, then the use of asymmetric kernels may provide better approximation results.

4 Conclusion

The optimization of Laguerre bases for the orthonormal series expansion of Volterra models has been addressed. A strict global optimal solution was derived by decomposing each multidimensional kernel of the model using a set of independent orthonormal bases, each of which parameterized by an individual Laguerre pole associated with the kernel dynamics along a particular dimension. This solution minimizes an upper bound of the error resulting from the truncation of the Laguerre series into a finite number of functions. It is an extension of the result in [5], where an analytic solution was obtained for the particular case in which a single Laguerre basis is used for expanding a given kernel along all its dimensions. It was proved that the particular and extended solutions are equivalent to each other when the Volterra kernels are symmetric.

Acknowledgements

The first and second authors acknowledge CNPq and Fapesp.

References


Appendix

Proof of Theorem 1

Following the lines of [5], it is possible to show that the terms \(S_{1,l} \) and \(S_{2,l} \) defined in (9) and (10), respectively, are related to the Laguerre poles and coefficients of the series expansion in (2), (3), and (5) as:

\[
\begin{align*}
S_{2,l} & = (1-p_{m,l})S_{1,l} + \sum_{l_{1}=1}^{\infty} \cdots \sum_{l_{m}=1}^{\infty} \left[ (p_{m,l}^{2} + (1-p_{m,l})\alpha_{l_{1},...l_{m}}^{2}) \right] (18)
\end{align*}
\]

In addition, using (2) and the orthonormality property of the Laguerre functions, it is straightforward to rewrite the squared norm \(\|h_{m}\|^2\) defined in (8) as:

\[
\begin{align*}
\|h_{m}\|^2 = \sum_{l_{1}=1}^{\infty} \cdots \sum_{l_{m}=1}^{\infty} \sum_{i_{1}=1}^{\alpha_{l_{1},...l_{m}}^{2}} i_{1} \cdots i_{m} = \sum_{l_{1}=1}^{\infty} \cdots \sum_{l_{m}=1}^{\infty} \left( (S_{1,l} - 1) p_{m,l}^{2} + (S_{2,l} - 2S_{1,l} + 1) p_{m,l} + S_{1,l} \right) \right)
\end{align*}
\]

(19)

Then, summing up (19) for \(l = 1, \ldots , m\) yields (7), which completes the proof. \(\Box\)