OPTIMAL POLE CONDITIONS FOR LAGUERRE MODELS THAT SATISFY SOME INTERPOLATION CONSTRAINTS, USING AN $\|\cdot\|_p$ NORM, $1 < p < \infty$

Tomás Oliveira e Silva

Departamento de Electrónica e Telecomunicações / IEETA
Universidade de Aveiro, 3810-193 Aveiro, Portugal
e-mail: `tos@det.ua.pt`

ABSTRACT

The optimal pole conditions for Laguerre models are available in the literature for the $\|\cdot\|_2$ norms (for continuous-time or discrete-time systems, with or without an impulsive input signal, and in the time or frequency domains). Recently, the author was able to extend the available results i) to other $\|\cdot\|_p$ norms, and ii) to models whose responses to known input signals satisfy, in the time and/or frequency domains, some interpolation constraints. In this paper we combine both extensions. It turns out that the optimality conditions for the poles of Laguerre models constrained as stated above, and using an $\|\cdot\|_p$ norm, have the same functional form as the already available optimality conditions: the last optimal weight of the model vanishes or the last optimal weight of the model of the next higher order vanishes.

1 Introduction

In linear system identification, linear prediction, linear echo cancellation, etc., there exist two possible choices of model structure. Either one uses a general IIR model structure, in which one can adjust the poles and zeros of the model, or one uses a linear-in-the-parameters model structure (depicted in figure 1), in which one fixes the poles of the model and one can adjust its zeros. The discrete-time FIR model falls in the second class of models. The main disadvantage of linear-in-the-parameters models over general IIR models is that their modeling capabilities are poorer, i.e., in general, they require more parameters to attain the same modelling performance. Its main advantage is that the optimization of its parameters is considerably simpler. It turns out that if one measures the performance of the model with the $\|\cdot\|_p$ norm, $1 < p < \infty$, of the error signal

$$e_m = y_m - y \tag{1}$$

of the model, where

$$y_m = \sum_{n=1}^{m} c_n\, x_n, \tag{2}$$

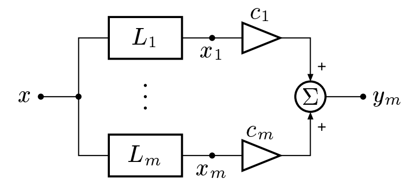

Figure 1: The block diagram of a linear-in-the-parameters model. This block diagram covers several distinct cases: discrete-time (DT) or continuous-time (CT) models, in the time or frequency domains. When the transfer functions L_n are the Laplace transforms of Laguerre functions, or z transforms of Laguerre sequences, one obtains a Laguerre model.

x_n being the so-called internal signals of the model and y its desired signal, then the performance surface is unimodal. This is a consequence of the fact that the Banach space associated with a standard $\|\cdot\|_p$ norm is strictly convex [1, 2]. When the c_n weights are optimal, the correspondent error signal will be denoted by ϵ_m.

The Laguerre model [3, 4, 5, 6, 7, 8] is another example of a linear-in-the-parameters model, in which the L_n transfer functions take the form

$$L_n(z) = \frac{\sqrt{1-u^2}}{z-u}\left(\frac{1-uz}{z-u}\right)^{n-1} \tag{3}$$

in the DT case, and the form

$$L_n(s) = \frac{\sqrt{-2u}}{s-u}\left(\frac{s+u}{s-u}\right)^{n-1} \tag{4}$$

in the CT case. The Laguerre models have one more parameter than the FIR models, which is the pole position u. It is precisely this extra parameter that makes the modelling capabilities of the Laguerre models potentially better than those of FIR models [7]. It makes sense to attempt to optimize the performance of the model also with respect to this parameter. This problem was addressed in the past for the $\|\cdot\|_2$ norm, when the input signal of the model is an impulse [9, 10, 11, 12], or an arbitrary signal [13, 14, 15]. Recently, the $\|\cdot\|_p$ norm case was solved [16], as was the case when the model satisfies some interpolation constraints [17]. In this paper we will combine the results of these two recently solved cases.

The rest of this paper is organized as follows. In section 2 we describe the interpolation constraints we can impose on the model. In section 3 we present, with a sketch of a proof, the optimality conditions for the pole of the Laguerre model. We conclude the paper with section 4, where an example is presented.

2 The interpolation constraints

Let

$$H_m(w) = \sum_{n=1}^{m} c_n L_n(w), \tag{5}$$

with $w = z$ (DT) or $w = s$ (CT), be the transfer function of the Laguerre model, and let $h_m(t)$ be the corresponding impulse response. The linear constraints we will impose on the weights of the Laguerre model will take of one of the forms

$$\begin{cases} \mathrm{Re}\big[H_m^{(k_i)}(w_i)\big] = \mathrm{Re}[h_i] \\ \mathrm{Im}\big[H_m^{(k_i)}(w_i)\big] = \mathrm{Im}[h_i], \end{cases} \tag{6}$$

or

$$h_m^{(k_i)}(t_i) = h_i, \tag{7}$$

where ${}^{(k_i)}$ denotes k_i times differentiation with respect to w or t, and $\mathrm{Re}[\cdot]$ and $\mathrm{Im}[\cdot]$ denote respectively the real and imaginary parts of a complex number. The m_0 real interpolation constraints are then of the form

$$\sum_{j=1}^{m} g_{i,j} c_j = h_i, \qquad i = 1, \ldots, m_0, \tag{8}$$

i.e., they are linear constraints, with

$$\begin{cases} g_{i,j} = \mathrm{Re}\big[L_j^{(k_i)}(w_i)\big] \\ g_{i,j} = \mathrm{Im}\big[L_j^{(k_i)}(w_i)\big], \end{cases} \tag{9}$$

or

$$g_{i,j} = l_j^{(k_i)}(t_i). \tag{10}$$

With these interpolations constraints it is possible, for example, to force the model to have a given DC gain, given zeros at given frequencies, a given relative order, given gain and phase margins at given frequencies, etc. It is also possible to impose interpolation constraints to $Y_m(w) = X(w)H_m(w)$ and to $y_m(t) = x(t) * h_m(t)$, or to their derivatives, when x is an input signal of a known form (a unit step, for example).

3 The Laguerre case

Due to severe lack of space, it is only possible to sketch the proof of the main result of the paper, which is presented at the very end of this section.

3.1 Preliminaries

It is possible to treat all cases (CT, DT, etc.) with a common mathematical notation. For that it is necessary to rewrite the $\|\cdot\|_p$ norms in the form

$$\|x\|_p^p = \langle |x|^p \rangle, \tag{11}$$

where $\langle\cdot\rangle$ depends on the Banach space being used and is defined in table 1 ($\mathtt{i}$ is the square root of -1). Let x be a complex function of the real variable α. It is possible to prove that

$$\frac{\partial |x|^p}{\partial \alpha} = p\,|x|^{p-2}\,\mathrm{Re}\left[x\left(\frac{\partial x}{\partial \alpha}\right)^*\right]. \tag{12}$$

3.2 The best way to deal with the interpolation constraints

One way to compute the optimal weights of an optimization problem with equality constraints is given by the limit

$$\lim_{\lambda\to\infty}\ \arg\min_{c_1,\ldots,c_m}\ \xi_m, \tag{13}$$

with

$$\xi_m = \|e_m\|_p^p + \lambda \sum_{i=1}^{m_0}\left(\sum_{j=1}^{m} g_{i,j}c_j - h_i\right)^2. \tag{14}$$

It can be shown that the weights are continuous and differentiable functions of λ, with finite limits when $\lambda \to \infty$. In [17] the case $p = 2$ is treated in detail.

3.3 The optimal weights

With the notation and results introduced in the previous two subsections, it is a simple but tedious task [16] to determine the system of equations that the optimal weights of the model must satisfy. These equations are, of course, obtained by equating $\partial\xi_m/\partial c_n$ to zero. This yields the "normal equations"

$$p\big\langle |\epsilon_m|^{p-2}\,\epsilon_m\,x_n^*\big\rangle + 2\lambda \sum_{i=1}^{m_0} g_{i,n} \sum_{j=1}^{m} (g_{i,j}c_j - h_i) = 0, \tag{15}$$

for $n = 1, \ldots, m$. Note that in these equations we have used the optimal error ϵ_m, because the weights are optimal.

Table 1: The definition of $\langle\cdot\rangle$ for several Banach spaces.

Type of signals	$\langle x \rangle$
DT, time domain	$\sum_{t=-\infty}^{+\infty} x(t)$
DT, frequency domain	$\frac{1}{2\pi}\int_{-\pi}^{+\pi} x(e^{\mathtt{i}\omega})\,d\omega$
CT, time domain	$\int_{-\infty}^{+\infty} x(t)\,dt$
CT, frequency domain	$\frac{1}{2\pi}\int_{-\infty}^{+\infty} x(\mathtt{i}\omega)\,d\omega$

3.4 The optimal pole conditions

In order to find the stationarity conditions with respect to the pole position of the Laguerre model it is necessary to compute the partial derivatives of the signals x_n with respect to u. In fact, according to (12) we have

$$\frac{\partial|\epsilon_m|^p}{\partial u} = p\,|\epsilon_m|^{p-2}\mathrm{Re}\left[\epsilon_m \sum_{n=1}^{m}\left(\frac{\partial c_n}{\partial u}x_n + c_n\frac{\partial x_n}{\partial u}\right)^*\right]. \tag{16}$$

Since these signals are obtained by filtering the model's input signal with stable linear time invariant filters, and since the input signal does not depend on the pole of the model, to compute these partial derivatives it is necessary to compute the partial derivatives of the transfer functions of these filters. These partial derivatives are given by [10, 12, 13]

$$\frac{\partial L_n(w)}{\partial u} = \frac{n\,L_{n+1}(w) - (n-1)L_{n-1}(w)}{\alpha}, \tag{17}$$

where $\alpha = 2u$ for the CT case, and $\alpha = 1-u^2$ for the DT case. It follows that

$$\frac{\partial x'_n}{\partial u} = \frac{n\,x'_{n+1} - (n-1)x'_{n-1}}{\alpha}. \tag{18}$$

Luckily, the partial derivatives of the $g_{i,j}$ with respect to u also follow formulas of the same kind. This is so because these coefficients are Laguerre functions, or their transforms, evaluated at certain points (w_i or t_i). The end result is

$$\frac{\partial g_{i,j}}{\partial u} = \frac{j\,g_{i,j+1} - (j-1)g_{i,j-1}}{\alpha}. \tag{19}$$

Using these formula to simplify $\partial\xi_m/\partial u = 0$, followed by an enforcement of the "normal equations" yields, after very tedious algebra, the conditions $c_m = 0$ and/or

$$p\langle|\epsilon_m|^{p-2}\,\epsilon_m\,x^*_{m+1}\rangle + 2\lambda\sum_{i=1}^{m_0} g_{i,m+1}\sum_{j=1}^{m}(g_{i,j}c_j - h_i) = 0. \tag{20}$$

This last condition together with (15) are the "normal equations" for a model with order $m+1$ in which *the last weight vanishes*. Since this result holds for any λ, and due to the smoothness of the variation of the optimal weights as a function of λ, the result will also be valid when $\lambda \to \infty$. Thus, the stationarity conditions we were looking for take the form

$$c_{m,m}\,c_{m+1,m+1} = 0, \tag{21}$$

where $c_{k,k}$ is the last optimal weight of a Laguerre model of order k.

4 Example

In this section we will attempt to approximate the output signal (y) of a system with transfer function [18]

$$H(z) = \sum_{k=1}^{5}\frac{r_k}{z - z_k} \tag{22}$$

Table 2: Poles and residues of $H(z)$.

k	z_k	r_k
1	$7/9$	$19/9$
2	$(1+\mathtt{i})/2$	$(4+7\mathtt{i})/5$
3	$(1-\mathtt{i})/2$	$(4-7\mathtt{i})/5$
4	$(4+7\mathtt{i})/13$	$(-8-10\mathtt{i})/13$
5	$(4-7\mathtt{i})/13$	$(-8+10\mathtt{i})/13$

by the output signal (y_m) of Laguerre models, when both systems are excited by a (causal) signal (x) with the z-transform

$$X(z) = z\,\frac{A}{z-0.6}. \tag{23}$$

The poles and their respective residues of $H(z)$ are presented in table 2. The amplitude (A) of the input signal is such that y has unit norm.

In our numerical experiment we illustrate what happens to the $\ell_5(\mathbb{N})$ norm of the optimal (with respect only to the weights) error signal of several Laguerre models, with orders $1,\ldots,10$, when we force the DC gain of the model to be equal to $H(1)$. Figure 2 presents the norm of the optimal error signal ϵ_m as a function of the pole position. Note that consecutive curves touch in extrema points of both curves. This is a consequence of the stationarity conditions, which state that in those points the last optimal weight of the higher order curve vanishes (the remaining common weights are the same, forcing the norms of the two optimal error signals to be equal). Usually, a local minimum of the order m curve touches a local maximum of the order $m+1$ curve.

In the first curve of figure 2, corresponding to $m = 1$, the sole weight of the model is enforced; recall that $H_m(1) = H(1)$. Thus, it is possible that $\|\epsilon_m\|_p$ may rise above 0 dB. Actually, this may happen to any curve. This happens in our example for values of u close to -1, because for those values of u the DC gain of the Laguerre transfer functions, which is $L_n(1) = \sqrt{(1+u)/(1-u)}$, is close to zero. This forces the weights of the model to take rather high values just to satisfy the DC gain constraint. If the enforced DC gain is far from its true value the result can be disastrous. We illustrate this in figure 3, where the condition $H_m(1) = 0.5\,H(1)$ was enforced.

References

[1] N. I. Achieser, *Theory of Approximation*, Dover Publications, Inc., New York, 1956, Published for the first time by Frederick Ungar Publishing Co. in 1956. Translated by Charles J. Hyman.

[2] P. J. Davis, *Interpolation and Approximation*, Dover Publications, Inc., New York, 1975, Published for the first time by Blaisdell Publishing Company in 1963.

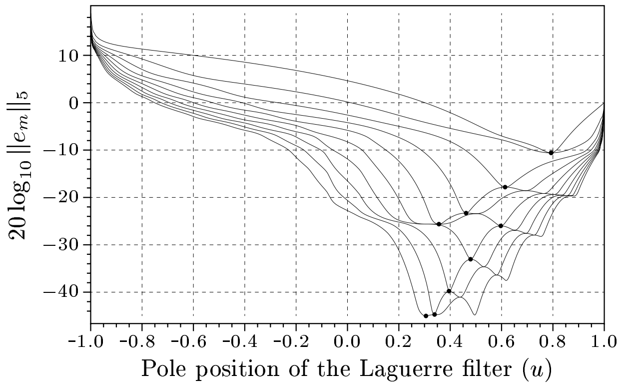

Figure 2: Graph of the $\ell_5(\mathbb{N})$ norm of the optimal error signal of the Laguerre models with orders $1, \ldots, 10$, constrained to have the same DC gain as $H(z)$.

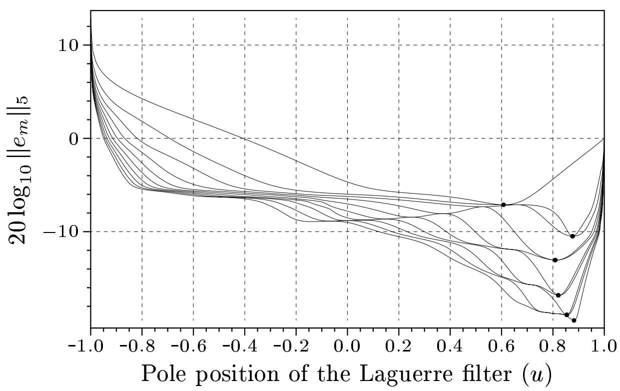

Figure 3: Same as figure 2 but with $H_m(1) = 0.5\,H(1)$.

[3] Y. W. Lee, "Synthesis of electrical networks by means of the Fourier transform of Laguerre's functions," *Journal of Mathematics and Physics*, vol. XI, pp. 83–113, 1931–1932.

[4] T. Kitamori, "Applications of orthogonal functions to the determination of process dynamic characteristics and to the construction of self-optimizing control systems," in *Proceedings of the First International Congress of the International Federation of Automatic Control*, Moscow, 1960, vol. II, pp. 613–617, Butterworths, London.

[5] T. L. Saaty and J. Bram, *Nonlinear Mathematics*, Dover Publications, Inc., New York, 1981, Published for the first time by McGraw-Hill Book Company, Inc., in 1964.

[6] R. E. King and P. N. Paraskevopoulos, "Digital Laguerre filters," *International Journal of Circuit Theory and Applications*, vol. 5, no. 1, pp. 81–91, Jan. 1977.

[7] B. Wahlberg, "System identification using Laguerre models," *IEEE Transactions on Automatic Control*, vol. 36, no. 5, pp. 551–562, May 1991.

[8] T. Oliveira e Silva, "Laguerre filters — an introduction," *Revista do DETUA*, vol. 1, no. 3, pp. 237–248, Jan. 1995.

[9] G. J. Clowes, "Choice of the time-scaling factor for linear system approximations using orthonormal Laguerre functions," *IEEE Transactions on Automatic Control*, vol. AC-10, no. 4, pp. 487–489, Oct. 1965.

[10] J. J. King and T. O'Canainn, "Optimum pole positions for Laguerre-function models," *Electronics Letters*, vol. 5, no. 23, pp. 601–602, Nov. 1969.

[11] M. A. Masnadi-Shirazi and N. Ahmed, "Optimal Laguerre networks for a class of discrete-time systems," *IEEE Transactions on Signal Processing*, vol. 39, no. 9, pp. 2104–2108, Sept. 1991.

[12] L. Wang and W. R. Cluett, "Optimal choice of time-scaling factor for linear system approximations using Laguerre models," *IEEE Transactions on Automatic Control*, vol. 39, no. 7, pp. 1463–1467, July 1994.

[13] T. Oliveira e Silva, "Optimality conditions for truncated Laguerre networks," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2528–2530, Sept. 1994.

[14] T. Oliveira e Silva, "Optimality conditions for Laguerre lattice filters," *IEEE Signal Processing Letters*, vol. 2, no. 5, pp. 97–98, May 1995.

[15] T. Oliveira e Silva, "On the determination of the optimal pole position of Laguerre filters," *IEEE Transactions on Signal Processing*, vol. 43, no. 9, pp. 2079–2087, Sept. 1995.

[16] T. Oliveira e Silva, "Optimal pole conditions for Laguerre and two-parameter Kautz models in several $\|\cdot\|_p$ norms. Part I: $1 < p < \infty$," Submitted to *IEEE Transactions on Signal Processing.*, Sept. 1999.

[17] T. Oliveira e Silva, "Optimal pole conditions for Laguerre and two-parameter Kautz models that satisfy some interpolation constraints in the time and/or frequency domains," Submitted to *IEEE Transactions on Automatic Control.*, Sept. 1999.

[18] T. Oliveira e Silva, "Stationary conditions for the L^2 error surface of the generalized orthonormal basis functions lattice filter," *Signal Processing*, vol. 56, no. 3, pp. 233–253, Feb. 1997.