Improved method for optimum choice of free parameter in orthogonal approximations

Noël Tanguy, Riwal Morvan, Pierre Vilbé, Léon-Claude Calvez

Improved Method for Optimum Choice of Free Parameter in Orthogonal Approximations

N. Tanguy, R. Morvan, P. Vilbé, and L. C. Calvez

Abstract—We report on our investigations to choose a free parameter to minimize the error energy when approximating a given signal with orthogonal basis functions. This method requires limited knowledge of the signal to be approximated and has a low computational cost.

Index Terms—Modeling, orthogonal functions, parameter estimation.

I. INTRODUCTION

Any arbitrary causal signal of finite energy can be expanded into an infinite series of complete orthogonal functions that often involve a free parameter. In theory, when an infinite number of terms is used in the expansion, the choice of the free parameter is somewhat arbitrary. In practice, however, a truncated series is used, and an immediate design problem consists in an optimal selection of the free parameter to minimize the truncation error. Such problems are often raised in model reduction and information compression. Scientists have investigated two different approaches. In the first one, the error energy is minimized; in the particular case of Laguerre functions, for example, Masnadi-Shirazi and Ahmed [1], [2] have developed an analytical approach that requires finding roots of possibly high-order polynomials. Reports have dealt with similar results obtained in an optimal pole position of Laguerre filters with arbitrary inputs [4] and [5]. Later, the optimality conditions for truncated Kautz series and networks involving two parameters have been presented [6], [7]. Such an approach generally entails an excessive computational cost. The alternative approach considered here will consist of minimizing an upper bound for the error energy. The case of continuous-time Laguerre functions has been previously studied in [8] and [9]. More recently, an intuitive suboptimal procedure to determine the poles of discrete-time Laguerre functions was presented [10], [11]. This last technique has the great merit that the derivation of a solution requires knowledge of only few numerical characteristics of the signal under consideration. A key point in the development of the proposed method is that several continuous- or discrete-time functions used in signal modeling satisfy a noteworthy equation. In [12] and [13], we considered a usual but particular form of it closely related to Laguerre, Charlier, and Meixner functions. Here, we will first state a generalized form of the noteworthy equation and, second, obtain a better upper bound for the quadratic error. Therefore, the purpose of the present correspondence is to show how the free parameter value is improved by a reiterated use of the noteworthy equation in the general case of orthogonal basis functions.

II. DEVELOPMENT

In this correspondence, we will consider a set of discrete-time or continuous-time real orthogonal functions $\varphi_n(t, \alpha)$ depending on a free parameter α to be optimally selected. The orthogonality of the

functions $\varphi_n(t, \alpha)$ can be noted $\langle \varphi_n, \varphi_m \rangle = \sigma_n^2(\alpha)\delta_{n,m}$, where $\sigma_n^2(\alpha)$ is the orthonormality constant, $\delta_{n,m}$ is the Kronecker symbol, and the inner product is defined according to the nature of the problem under consideration through

$$\langle f, g \rangle \triangleq \int_{t_0}^{t_I} \varpi(t)f(t)g(t)\, dt$$

(continuous case)

$$\langle f, g \rangle \triangleq \sum_{i=0}^{I} \varpi(t_i)f(t_i)g(t_i) \quad \text{(discrete case)}$$

where $t_i (i = 0, 1, \cdots, I)$ denotes discrete points in $[t_0, t_I]$ and $\varpi(t)$ a non-negative weighting function.

Let us suppose that $f(t)$ is a well-behaved real-valued signal, which can be represented by the infinite expansion as

$$f(t) = \sum_{n=0}^{\infty} c_n(\alpha)\varphi_n(t, \alpha), \quad c_n(\alpha) = \langle f, \varphi_n \rangle / \sigma_n^2(\alpha) \tag{1}$$

where $c_n(\alpha)$ results from the classical theory of orthogonal expansions.

Let us consider the practical truncated series $\tilde{f}(t) = \sum_{n=0}^{N-1} c_n(\alpha)\varphi_n(t, \alpha)$, which, for a given α, is known to be the best N-terms approximation to $f(t)$ in the sense of minimizing the relative weighted quadratic error $q_N(\alpha) \triangleq \|\tilde{f} - f\|^2 / \|f\|^2 = \sum_{n=N}^{\infty} \sigma_n^2(\alpha)c_n^2(\alpha)/\|f\|^2$, where $\|f\|^2 \triangleq \langle f, f \rangle$. We should note that $q_N(\alpha)$ depends on α. It is then clear that the quadratic error can be reduced by properly choosing α.

A key point in this development is that several continuous- or discrete-time orthogonal functions satisfy the noteworthy equation

$$\mathcal{L}_\alpha \varphi_n(t, \alpha) = \lambda(n)\varphi_n(t, \alpha) \tag{2}$$

where $\lambda(n)$ is a non-negative nondecreasing sequence independent of time, and $\mathcal{L}_\alpha$ denotes a linear operator on time t, which can be put into the following form:

$$\mathcal{L}_\alpha f(t) = \sum_{j=0}^{J} h_j(t, \alpha)\mathcal{H}_{j,\alpha} f(t) \tag{3}$$

where the coefficients $h_j(t, \alpha)$ depend on t and α but are independent of n. $\mathcal{H}_{j,\alpha}$ denotes linear operators, i.e., delay, derivative, or integral operators, operating on time.

When assuming that $\varphi_n(t, \alpha)$ satisfies (2), applying the linear operator $\mathcal{L}_\alpha$ to (1) and using the linearity property of the inner product, which is assumed to exist, lead to the relation

$$\langle f, \mathcal{L}_\alpha f \rangle = \sum_{n=0}^{\infty} \lambda(n)\sigma_n^2(\alpha)c_n^2(\alpha)$$

$$\geq \lambda(N) \sum_{n=N}^{\infty} \sigma_n^2(\alpha)c_n^2(\alpha) \tag{4}$$

which gives an upper bound for $q_N(\alpha)$

$$q_N(\alpha) \leq F(\alpha)/\lambda(N) \quad \text{for } \lambda(N) \geq F(\alpha) \tag{5}$$

where

$$F(\alpha) \triangleq \langle f, \mathcal{L}_\alpha f \rangle / \|f\|^2 \tag{6}$$

One should note that as $q_N(\alpha)$ cannot be greater than unity, the condition on $\lambda(N)$ is added in (5). Therefore, when the specification

of the signal $f(t)$ is limited to the knowledge of $F(\alpha)$, inequality (5) implies that the best choice for α is the value α_0 that minimizes $F(\alpha)$, and the maximum error in $f(t)$ approximation is then $F(\alpha_0)/\lambda(N)$. It is worth noting that the optimal parameter α_0 is independent of the number of basis functions to be used.

Example: Let us consider the functions

$$\varphi_n(t, \alpha) \triangleq e^{-(\alpha^2 t^2/2)} H_n(\alpha t) \tag{7}$$

where $H_n(t)$ denotes the Hermite polynomials [14], and α is a strictly positive time-scale parameter to be optimally selected. The functions $\varphi_n(t, \alpha)$ are orthogonal on $]-\infty, +\infty[$ with the weight function $\varpi(t) = 1$

$$\langle \varphi_n, \varphi_m \rangle = \frac{2^n n! \sqrt{\pi}}{\alpha} \delta_{n,m}.$$

Using the well-known differential equation verified by the Hermite polynomials $y = H_n(t)$:

$$y'' - 2ty' + 2ny = 0$$

with $y' \triangleq dy/dt$, it is a standard exercise to show that the following relation holds:

$$-\frac{1}{\alpha^2} \varphi_n''(t, \alpha) + (\alpha t)^2 \varphi_n(t, \alpha) = (2n + 1) \varphi_n(t, \alpha). \tag{8}$$

Thus, defining the operator $\mathcal{L}_\alpha$ as

$$\mathcal{L}_\alpha f(t) \triangleq -\frac{1}{\alpha^2} f''(t) + (\alpha t)^2 f(t) \tag{9}$$

yields

$$\mathcal{L}_\alpha \varphi_n(t, \alpha) = \lambda(n) \varphi_n(t, \alpha). \quad \lambda(n) = 2n + 1.$$

A well-behaved real-valued function $f(t)$ of finite energy being given, using (9), (6), and $\int_{-\infty}^{+\infty} ff'' \, dt = -\int_{-\infty}^{+\infty} f'^2 \, dt$ leads to

$$F(\alpha) = \frac{1}{\alpha^2} \frac{\|f'\|^2}{\|f\|^2} + \alpha^2 \frac{\|tf\|^2}{\|f\|^2}.$$

Thus, in the case of approximation with Hermite functions, the optimal α-value minimizing $F(\alpha)$ and, therefore, the upper bound for $q_N(\alpha)$, is

$$\alpha_0 = \sqrt{\frac{\|f'\|}{\|tf\|}}. \tag{10}$$

Note that the optimal value α_0 given by (10) always exists and is very simple to compute.

Remark: $\mathcal{L}_\alpha$ has no real meaning. Nevertheless, many orthogonal functions satisfy a well-known differential equation (in continuous-time case) or a recurrence relation (in discrete-time case) from which $\mathcal{L}_\alpha$ can be readily derived [14], [15].

III. REITERATED USE OF THE OPERATOR

With a view to obtain a better upper bound for the quadratic error, the proposed method may be improved by applying the $\mathcal{L}_\alpha$ operator m times to the functions $\varphi_n(t, \alpha)$; therefore, we get

$$\mathcal{L}_\alpha^m \varphi_n(t, \alpha) = \lambda^m(n) \varphi_n(t, \alpha) \tag{11}$$

which yields the new relation

$$q_N(\alpha) \leq F_m(\alpha)/\lambda^m(N) \quad \text{for } \lambda^m(N) \geq F_m(\alpha) \tag{12}$$

with

$$F_m(\alpha) \triangleq \langle f, \mathcal{L}_\alpha^m f \rangle / \|f\|^2. \tag{13}$$

This technique increases the exponent of $\lambda(N)$ in (12), and thus, for N great enough, it yields a better upper bound for the quadratic error.

Remark: N can be selected to guarantee $q_N(\alpha) \leq \varepsilon$. In this way, we first calculate α_0, which is independent of N, and second, we determine N so that $F_m(\alpha_0)/\lambda^m(N) \leq \varepsilon$.

IV. ILLUSTRATIVE EXAMPLE

In order to illustrate the method, the discrete Laguerre functions will be used as basis functions. The first point consists of determining an analytic expression for the upper bound for the quadratic error using characteristic components of the function to be modeled.

Let $\varphi_n(k, a) \triangleq b^k l_n(k, a)$ be the weighted discrete Laguerre functions, where $l_n(k, a)$ stands for the Laguerre functions used in [1]–[3]. Functions $\varphi_n(k, a)$ are orthonormal on $[0, \infty[$ with respect to the weight function $\varpi(k) = b^{-2k}$, and a is a free parameter to be optimized. Using, for example, the Z transform

$$\sum_{k=0}^{\infty} \varphi_n(k, a) z^{-k} = \frac{z\sqrt{1-a^2}}{z - ab} \left(\frac{b - az}{z - ab}\right)^n$$

it is a standard exercise to show that the following relation holds:

$$\mathcal{L}_a \varphi_n(k, a) = n \varphi_n(k, a) \tag{14}$$

where

$$\mathcal{L}_a f(k) = \frac{-1}{b(1 - a^2)} \{(k + 1)af(k + 1) - [k(1 + a^2) + a^2]bf(k) + kab^2 f(k - 1)\} \tag{15}$$

according to (3) with $t = k$ (discrete time). The use of expression (15) for $\mathcal{L}_a f(k)$ yields a first good method to choose the free parameter a [13]. In some cases, however, providing a better upper bound for the quadratic error is important. Thus, the operator $\mathcal{L}_a^2$ can be used to improve the method described. Using (14) leads to

$$\mathcal{L}_a^2 \varphi_n(k, a) = n^2 \varphi_n(k, a).$$

Therefore, with $m = 2$ in (13), it can be easily shown that the following relation holds:

$$F_2(a) = \frac{1}{(1 - a^2)^2} [(m_3 + 2m_1 + 1)a^4 - (4m_4 + 6m_2)a^3 + (2m_5 + 4m_3 + 4m_1 + 1)a^2 - (4m_4 + 2m_2)a + m_3]$$

with

$$m_1 \triangleq \sum_{k=0}^{\infty} kb^{-2k} f^2(k)/\|f\|^2$$

$$m_2 \triangleq \sum_{k=0}^{\infty} kb^{-(2k-1)} f(k-1)f(k)/\|f\|^2$$

$$m_3 \triangleq \sum_{k=0}^{\infty} k^2 b^{-2k} f^2(k)/\|f\|^2$$

$$m_4 \triangleq \sum_{k=0}^{\infty} k(k-1)b^{-(2k-1)} f(k-1)f(k)/\|f\|^2$$

$$m_5 \triangleq \sum_{k=0}^{\infty} k(k-1)b^{-(2k-2)} f(k-2)f(k)/\|f\|^2$$

denoting "moments." To determine the optimal parameter a, setting the derivative of $F(a)$ with respect to a equal to 0, and simplifying the resulting equation gives

$$(2m_4 + 3m_2)a^4 - (2m_5 + 6m_3 + 8m_1 + 3)a^3 + (12m_4 + 12m_2)a^2 - (2m_5 + 6m_3 + 4m_1 + 1)a + 2m_4 + m_2 = 0. \tag{16}$$
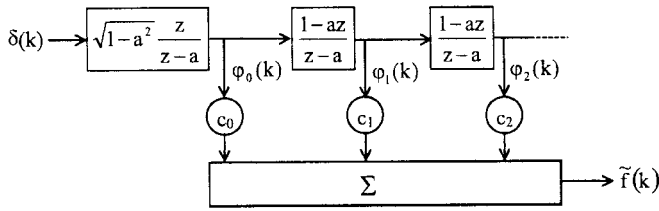
Fig. 1.   Discrete Laguerre model with $b = 1$.

Remark: Note that the term $1/(1 - a^2)$ in (15) implies that for discrete-time Laguerre functions, $F_m(a) \triangleq \langle f, \mathcal{L}_a^m f \rangle / \|f\|^2$ will be always in the form $F_m(a) = P_{2m}(a)/(1 - a^2)^m$, where $P_{2m}(a)$ is a $2m$-degree polynomial. Because $-1 < a < 1$ and (12), $F_m(a)$ is always positive or equal to zero, and $\lim_{a \to \pm 1} F_m(a) = +\infty$ (except for $f[k] \equiv 0$). Therefore, there is always a value in $]-1, 1[$ for the parameter a that minimizes $F_m(a)$.

V. NUMERICAL EXAMPLE

The stated problem consists of determining the optimum eighth-order discrete Laguerre model (Fig. 1 using $b = 1$) for an "unknown" plant, which here is simulated by the discrete rational transfer function corresponding to [3, Example 3.2]:

$$F(z) = \frac{z(z^3 - 3.8z^2 + 1.96z - 0.297)}{(z - 0.2)(z - 0.3)(z + 0.5)(z - 0.9)}.$$

All calculations will be based only on the knowledge of the samples of $F(z)$ impulse response. The optimal value for the parameter a computed in [3] is $a_0 = 0.4305$. The quadratic error is reduced to $q = 0.0149$, which can be compared with $q = 0.1071$ obtained with the arbitrary value $a = 0.7$. This result shows the importance of optimizing the free parameter in Laguerre modeling. Whereas the above optimization requires root-finding of polynomial equations of degrees 31 and 35, our technique reduces the amount of calculus drastically while retaining a good optimization. Using $b = 1$ enables us to compute the moments from the numerical values of the impulse response $f(k)$; they are obtained as $\|f\|^2 = 16.6454$, $m_1 = 3.4002$, $m_2 = 2.9155$, $m_3 = 29.2896$, $m_4 = 28.3785$, $m_5 = 32.3523$, which lead to the roots of (16): 2.3007, $0.6992 \pm 0.6530i$, 0.4326. Therefore, the proposed method yields $a_0' = 0.4326$. This value, obtained via the resolution of a degree-4 polynomial equation, has to be compared with the optimal value $a_0 = 0.4305$, whereas we previously [13] had $a_0 = 0.4492$ with $m = 1$ in (13).

VI. CONCLUSION

This correspondence presents a general technique for choosing a free parameter in orthogonal approximations. It is divided into two steps. The first one provides an analytical expression for the upper bound for the quadratic error; the obtained expression only depends on the free parameter of the basis functions. The second is the applicative phase and consists of minimizing this analytical expression in order to optimize the free parameter of the basis functions. Analytical results in Hermite and Laguerre cases are derived, and a numerical example is presented. We showed that the proposed procedure requires much less calculus than an optimal method and can provide a result very close to the optimal one. We simply and successfully implemented this technique on fixed-point processors.

REFERENCES

[1] M. A. Masnadi-Shirazi and N. Ahmed, "Laguerre approximation of nonrecursive discrete-time systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Albuquerque, NM, Apr. 3–6, 1990, pp. 1309–1312.

[2] ——, "Optimum Laguerre networks for a class of discrete-time systems," *IEEE Trans. Signal Processing*, vol. 39, pp. 2104–2108, 1991.

[3] M. A. Masnadi-Shirazi, "Optimum synthesis of linear discrete-time systems using orthogonal Laguerre sequences," Ph.D. dissertation, Univ. New Mexico, Albuquerque, May 1990.

[4] T. Oliveira e Silva, "Optimality conditions for truncated Laguerre networks," *IEEE Trans. Signal Processing*, vol. 42, pp. 2528–2530, Sept. 1994.

[5] ——, "On the determination of the optimal pole position of Laguerre filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 2079–2087, Sept. 1995.

[6] ——, "Optimality conditions for truncated Kautz networks with two periodically repeating complex conjugate poles," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 342–346, Feb. 1995.

[7] A. C. Den Brinker, F. P. A. Benders, and T. Oliveira e Silva, "Optimality conditions for truncated Kautz series," *IEEE Trans. Circuits Syst. II*, vol. 43, pp. 117–122, Feb. 1996.

[8] G. J. Clowes, "Choice of the time-scaling factor for linear system approximations using orthonormal Laguerre functions," *IEEE Trans. Automat. Contr.*, vol. AC-10, pp. 487–489, Oct. 1965.

[9] T. W. Parks, "Choice of time scale in Laguerre approximations using signal measurements," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 511–513, Oct. 1971.

[10] Y. Fu and G.A. Dumont, "An optimum time scale for discrete Laguerre network," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 934–938, June 1993.

[11] ——, "On determination of Laguerre filter pole through step or impulse response data," in *Proc. IFAC 12th Triennal World Cong.*, Sydney, Australia, 1993, pp. 35–39.

[12] N. Tanguy, "La transformation de Laguerre discrète," Thèse de doctorat, Univ. Brest, Brest, France, Dec. 1994.

[13] N. Tanguy, P. Vilbé, and L. C. Calvez, "Optimum choice of free parameter in orthonormal approximations," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1811–1813, 1995.

[14] P. Beckmann, *Orthogonal Polynomials for Engineers and Physicists*. Boulder, CO: Golem, 1973.

[15] G. Szegö, *Orthogonal Polynomials*. Providence, RI: Amer. Math. Soc., 3rd ed., 1967.