

Optimum choice of free parameter in orthonormal approximations

Noël Tanguy, Pierre Vilbé, Léon-Claude Calvez

► **To cite this version:**

Noël Tanguy, Pierre Vilbé, Léon-Claude Calvez. Optimum choice of free parameter in orthonormal approximations. IEEE Transactions on Automatic Control, Institute of Electrical and Electronics Engineers, 1995, 40 (10), pp.1811-1813. <10.1109/9.467666>. <hal-00439477>

HAL Id: hal-00439477

<http://hal.univ-brest.fr/hal-00439477>

Submitted on 7 Dec 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimum Choice of Free Parameter in Orthonormal Approximations

N. Tanguy, P. Vilbé, and L. C. Calvez

Abstract—This correspondence investigates the choice of a free parameter, usually related to time scale, that minimizes the error energy when approximating a given signal with some widely used orthonormal basis functions. The proposed solution is the best that can be achieved with the limited required knowledge of the signal. It is appealing for experimental data for which no exact mathematical expression is available.

I. INTRODUCTION

In recent years there has been renewed interest in the use of orthogonal functions in approximation, modeling, filtering, identification, and control. Particularly, the discrete Laguerre functions, a complete orthogonal set on $\ell^2[0, \infty)$ have been considered often. Any arbitrary causal signal of finite energy, e.g., an impulse response, can be expanded into an infinite series of orthogonal Laguerre functions which involve a free parameter, closely related to the time scale. In theory, when infinitely many terms are used in the expansion, the choice of the parameter is somewhat arbitrary. In practice, however, a truncated series is used, and an immediate design problem is to optimally select the free parameter so as to minimize the truncation error. With a view to minimizing the error energy, Masnadi-Shirazi and Ahmed [2], [3] have derived an analytical approach which requires finding real roots of possibly high order polynomials, lying in the interval $]0, 1[$. For numerical approaches, see references given in [3].

In a recent publication, an alternative way has been presented by Fu and Dumont [1]. The idea is to minimize a performance index J defined as a linearly weighted sum of squared Laguerre coefficients. Intuitively, enforced convergence rate is then expected. A great merit of this technique is that it yields an analytical solution, requiring only a few numerical characteristics of the signal under consideration. The purpose of the present correspondence is twofold: first, to show that if the number of basis functions to be used is chosen large enough, then the performance index J proposed in [1] is related to an upper bound of the error energy, and second, to offer a general method which applies to a wide class of continuous or discrete time orthonormal basis functions.

II. DEVELOPMENT

To achieve a unified treatment, valid for both continuous- and discrete-time signals, inner product notation will be used. Let $\varpi(t)$ denote a nonnegative weighting function of the real variable t defined over some finite or infinite interval $[t_0, t_I]$. Given two well-behaved real-valued function $f(t)$ and $g(t)$ of the real variable t defined over $[t_0, t_I]$, it is assumed that their inner product is defined, according to the nature of the problem under consideration, through the following integral

$$\langle f, g \rangle \doteq \int_{t_0}^{t_I} \varpi(t) f(t) g(t) dt \quad (1a)$$

Manuscript received August 2, 1994; revised March 14, 1995.

The authors are with Laboratoire d'Electronique et Systèmes de Télécommunications, L.E.S.T. U.R.A. C.N.R.S. no. 1329, Université de Bretagne Occidentale (U.B.O.), Brest, France.

IEEE Log Number 914041.

or through the following discrete sum

$$\langle f, g \rangle \doteq \sum_{i=0}^I \varpi(t_i) f(t_i) g(t_i) \quad (1b)$$

where $t_i (i = 0, 1, \dots, I)$ denote discrete points in $[t_0, t_I]$.

Since any orthogonal system of functions can be readily normalized by dividing each member of the set by its norm, we shall consider orthonormal functions $\varphi_n(t, \alpha)$ satisfying

$$\langle \varphi_n, \varphi_m \rangle = \delta_{n,m} \quad (2)$$

where α denotes a free parameter to be optimally selected. This parameter may be a scale factor as in continuous-time Hermite functions $\mathcal{H}_n(\alpha t)$ or some real number as in discrete-time Charlier functions $\mathcal{C}_n(t, \alpha)$.

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

$$f(t) = \sum_{n=0}^{\infty} c_n(\alpha) \varphi_n(t, \alpha) \quad (3)$$

where, from the classical theory of orthonormal expansions

$$c_n(\alpha) = \langle f, \varphi_n \rangle. \quad (4)$$

Consider the practical truncated series

$$\tilde{f}(t) \doteq \sum_{n=0}^{N-1} c_n(\alpha) \varphi_n(t, \alpha) \quad (5)$$

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 $q_N(\alpha) \doteq \|\tilde{f} - f\|^2 / \|f\|^2$. Using $\|f\|^2 \doteq \langle f, f \rangle$ and the orthonormality property (2) it is a standard exercise to show that the following holds

$$q_N(\alpha) \doteq \sum_{n=N}^{\infty} c_n^2(\alpha) / \|f\|^2 = 1 - \sum_{n=0}^{N-1} c_n^2(\alpha) / \|f\|^2. \quad (6)$$

Usually, the quadratic error can be reduced further by a proper choice of the free parameter α as shown in the following.

A key point in the development of the proposed method is the observation that several continuous- or discrete-time orthonormal functions widely used in signal modeling and related to the so-called classical orthogonal polynomials satisfy the noteworthy equation [4], [5]

$$\mathcal{L}_\alpha \varphi_n(t, \alpha) = \lambda(n) \varphi_n(t, \alpha) \quad (7)$$

with \mathcal{L}_α denoting a linear operator defined as

$$\mathcal{L}_\alpha f(t) \doteq A(t, \alpha) \mathcal{D}^2 f(t) + B(t, \alpha) \mathcal{D}^1 f(t) + C(t, \alpha) \mathcal{D}^0 f(t) \quad (8)$$

where the coefficients A, B , and C which are characteristic of the set $\{\varphi_n(t, \alpha)\}$ depend on the time and the parameter α but are independent of n . In the context of continuous-time signals, $\mathcal{D}^i f(t)$ stands for the i th derivative $d^i f/dt^i$ while in the context of discrete-time signals, $\mathcal{D}^i f(t) \doteq f(t+i-1)$. As for the $\lambda(n)$ they are independent of t and form a sequence of increasing numbers

$$\lambda(0) = 0 < \lambda(1) < \lambda(2) < \lambda(3) < \dots$$

Until further notice, we assume $\varphi_n(t, \alpha)$ satisfies (7). With this equation in mind, applying the linear operator \mathcal{L}_α to (3) yields

$$\mathcal{L}_\alpha f(t) = \sum_{n=0}^{\infty} \lambda(n) c_n(\alpha) \varphi_n(t, \alpha)$$

now, using the linearity property of the inner product which is assumed to exist, on account of (4), we obtain

$$\langle f, \mathcal{L}_\alpha f \rangle = \sum_{n=0}^{\infty} \lambda(n) c_n^2(\alpha) \geq \lambda(N) \sum_{n=N}^{\infty} c_n^2(\alpha) \quad (9)$$

which gives an upper bound for $q_N(\alpha)$ in (6)

$$q_N(\alpha) \leq F(\alpha) / \lambda(N) \quad (10)$$

where $F(\alpha)$ is a mathematical expression defined by

$$F(\alpha) \doteq \langle f, \mathcal{L}_\alpha f \rangle / \|f\|^2 \quad (11)$$

which bears some information about the signal f . When this signal is known via experimental data, $F(\alpha)$ can be readily determined by computer.

Since $q_N(\alpha)$ cannot be greater than unity, notice that to get a useful bound, N must be great enough to ensure $\lambda(N) > F(\alpha)$.

Denoting by \mathcal{C}_F the class of signals with given $F(\alpha)$, it is a simple matter to prove the existence of a signal $f \in \mathcal{C}_F$ that achieves the bound in (10). Assuming $0 < F(\alpha) < \lambda(N)$, let

$$f(t) = K[\sqrt{1 - F(\alpha)/\lambda(N)} \varphi_0(t, \alpha) + \sqrt{F(\alpha)/\lambda(N)} \varphi_N(t, \alpha)]$$

where K is independent of time. Using (9)

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

therefore f is in class \mathcal{C}_F . Using (6), $q_N(\alpha) = F(\alpha)/\lambda(N)$ and consequently the bound in (10) is the best that can be achieved when the specification of the signal f is limited to the knowledge of $F(\alpha)$. In this case, the best choice of the parameter α is that value α_0 which minimizes $F(\alpha)$ and the maximum error over the class of all signals in \mathcal{C}_F is then $F(\alpha_0)/\lambda(N)$. It is worth noting that the optimal parameter α_0 is independent of the number N of basis functions to be used. Thus, α_0 can be computed in a first time and N can be chosen afterwards: for instance, one can choose N such that the upper bound $F(\alpha_0)/\lambda(N)$ is small enough or such that the exact $q_N(\alpha_0)$ in (6) is small enough.

Remark 1: Frequently, we have $\lambda(n) = n$; this is, for example, the case for Hermite, Laguerre, Charlier, and Meixner functions.

Remark 2: Because $\lambda(n) = n$, the performance index J proposed in [1] in the particular case of discrete Laguerre functions expansions is seen from (9) and (11) to be related to $F(\alpha)$ by $\|f\|^2[F(\alpha) + 1]$, thus minimizing J is equivalent to minimizing the maximum error encountered for the equivalence class of signals with given $F(\alpha)$.

Remark 3: By considering the particular case of continuous time Laguerre functions $\varphi_n(t, \alpha) = \sqrt{\alpha} \ell_n(\alpha t)$ and the appropriate \mathcal{L}_α one obtains Parks' method [6]; notice that this method can be somewhat extended by considering exponentially weighted Laguerre functions.

III. ILLUSTRATIVE EXAMPLE

Let $\varphi_n(k, a) \doteq b^k \ell_n(k, a)$ denote weighted discrete Laguerre functions where $\ell_n(k, a)$ stand for the Laguerre functions used in [2], [3], and [7]; the definition of [1] differs by one unit delay. Functions

$\varphi_n(k, a)$ are orthonormal on $[0, \infty)$ with respect to the weighting function $\varpi(k) = b^{-2k}$. The use of weighted Laguerre functions ($b > 1$) has the advantage that it may achieve a better approximation to the initial transient response than nonweighted functions ($b = 1$).

Problem Statement: Given $b \geq 1$ and a discrete signal $f(k)$ with known "moments"

$$m_1 \doteq \sum_{k=1}^{\infty} k[b^{-k}f(k)]^2/\|f\|^2$$

$$m_2 \doteq \sum_{k=1}^{\infty} k[b^{-(k-1)}f(k-1)][b^{-k}f(k)]/\|f\|^2$$

where

$$\|f\|^2 \doteq \sum_{k=0}^{\infty} [b^{-k}f(k)]^2$$

find the free parameter $\alpha = a$ such that the weighted sum of the squares of the errors is least.

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 holds

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

where $\mathcal{L}_a f(k) = -\frac{(k+1)a}{(1-a^2)b}f(k+1) + \frac{k(1+a^2)+a^2}{1-a^2}f(k) - \frac{kab}{1-a^2}f(k-1)$ which is in the form of (8) with $t = k$ (discrete time). Using the last expression for $\mathcal{L}_a f$ and definition (11) one obtains

$$F(a) = \frac{(1+m_1)a^2 - 2m_2a + m_1}{1-a^2}$$

Let $\rho \doteq (2m_1 + 1)/(2m_2)$ for which it can be shown that $|\rho| \geq 1$ is always true. The optimal value of a minimizing $F(a)$ is then

$$a_0 = \rho - \sqrt{\rho^2 - 1}$$

which results in a Laguerre parameter that minimizes the maximum of the normalized mean square error over the class of all signals with given $F(a)$, i.e., with given m_1 and m_2 . The minimum of $F(a)$ may be written

$$F(a_0) = \frac{[(1+m_1)\rho - m_2]a_0 - 0.5}{1 - \rho a_0}$$

and provided $N \geq F(a_0)$, for any signal with given m_1 and m_2 , the relative weighted quadratic error is $q_N(a_0) \leq F(a_0)/N$. This last bound is the tightest bound derivable with knowledge of the signal limited to m_1 and m_2 .

Remark: Moments m_1 and m_2 defined above are preferred to M_1 and M_2 of [1] which bear identical information but require the superfluous extra evaluation of $\Delta f(k) = f(k+1) - f(k)$ for each k .

Numerical Example: Consider the discrete signal [7, Example 3.2]

$$f(k) = -(0.2)^k + 0.5(0.3)^k + 3(-0.5)^k - 1.5(0.9)^k$$

Using $b = 1$, let the moments be computed from the numerical values of $f(k)$; these are obtained as $m_1 = 3.40103$ and $m_2 = 2.91636$ which yields $a_0 = 0.4492$ to be compared with $a'_0 = 0.4305$ which is the optimal value computed in [7], for an eight-terms Laguerre approximation, via solving polynomial equations of degree 31 and 35 for roots in $]0, 1[$ and subsequently retaining that root which gives the smallest value for the error energy.

Using the method in [1] yields $a_0 = 0.4492$ as above which is not surprising since, in the particular case of Laguerre expansion with $b = 1$, minimizing J is equivalent to minimizing $F(\alpha)$ (see Remark 2).

IV. CONCLUSION

With a view to selecting a free parameter in orthonormal approximations, a quite general and easy-to-use procedure has been proposed. It does not require complete specification of the signal to be approximated and is thus suited to problems dealing with experimental data. The resulting parameter is the best that can be achieved with the limited required knowledge of the signal.

REFERENCES

- [1] 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.
- [2] M. A. Masnadi-Shirazi and N. Ahmed, "Laguerre approximation of nonrecursive discrete-time systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing* Albuquerque, NM, Apr. 3-6, 1990, pp. 1309-1312.
- [3] —, "Optimum Laguerre networks for a class of discrete-time systems," *IEEE Trans. Signal Processing*, vol. 39, pp. 2104-2108, Sep. 1991.
- [4] G. Szegő, *Orthogonal Polynomials* 3rd ed. (Colloquium Publications) vol. 23. Providence, RI: American Math. Society, 1967.
- [5] P. Beckmann, *Orthogonal Polynomials for Engineers and Physicists*. Boulder, CO: Golem, 1973.
- [6] T. W. Parks, "Choice of time scale in Laguerre approximations using signal measurements," *IEEE Trans. Automat. Contr.*, vol. 16, pp. 511-513, Oct. 1971.
- [7] M. A. Masnadi-Shirazi, "Optimum synthesis of linear discrete-time systems using orthogonal Laguerre sequences," Ph.D. dissertation, Univ. New Mexico, May 1990.