On the Determination of the Optimal Center and Scale Factor for Truncated Hermite Series

T. Oliveira e Silva*
Universidade de Aveiro / INESC Aveiro
Portugal
tos@inesca.pt

H. J. W. Belt
Eindhoven University of Technology
The Netherlands
H.J.W.Belt@ele.tue.nl

ABSTRACT

Signals that are fairly concentrated in time or space can often be conveniently described by a truncated Hermite series. The rate of convergence of such series depends on the center and scale factor of the Hermite functions. In this paper we present some results concerning the determination of the optimal values of these two important parameters. We address the problem of the approximation of one-dimensional functions defined in the continuous interval $(-\infty, \infty)$.

1 Introduction

In some signal and image processing applications it is necessary to approximate signals which either have a finite support or are very concentrated (in time or space). One way to approximate this type of signals is to use a truncated Hermite series, since the factor $\exp(-x^2/2)$ in the Hermite functions approaches zero rather quickly when $|x|$ is substantially larger that 1. There is, however, one problem with this approach: it is necessary to find a good center and a good scale factor for the Hermite functions in order to have fast convergence of the series. This is the point we will address in this paper. We note that truncated Hermite expansions were used recently in the approximation of biomedical signals [1]. Other applications of Hermite series can be found in [2], [3].

The results of this paper were inspired by similar results for truncated Laguerre expansions [4], [5], [6], [7]. Due to lack of space we will discuss the approximation of one-dimensional signals only.

The structure of this paper is the following. Section 2 is concerned with the Hermite functions, and with some of their properties. Section 3 presents the stationarity conditions of the squared error of a truncated Hermite series. In Section 4 we describe a method to find a "good" center and a "good" scale factor for the Hermite functions based on measurements of the function to be approximated. Lastly, in Section 5 we present a simple example of the main results of the paper.

* This author gratefully acknowledges the financial support of JNICT, Junta Nacional de Investigação Científica e Tecnológica, Portugal.

2 The Hermite polynomials, the Hermite functions, and some of their properties

The Hermite polynomials form one of the families of classical orthogonal polynomials and may be defined for $n > 0$ by the recurrence relation [8]

$$H_{n+1}(x) = 2x\, H_n(x) - 2n\, H_{n-1}(x) \tag{1}$$

with $H_0(x) = 1$. Another useful recurrence relation is

$$H_{n+2}(x) = [4x^2 - 2(2n+1)]\, H_n(x) - 4n(n-1)\, H_{n-2}(x) \tag{2}$$

with $H_0(x) = 1$ and $H_1(x) = 2x$, and for $n \geq 0$. It is also easy to prove that

$$H_n'(x) = 2n\, H_{n-1}(x). \tag{3}$$

It is well known that the Hermite polynomials obey the orthogonality condition

$$\int_{-\infty}^{\infty} H_m(x) H_n(x)\, e^{-x^2}\, dx = 2^n n! \sqrt{\pi}\, \delta_{mn}, \quad m, n \geq 0.$$

From the Hermite polynomials, which are orthogonal w.r.t. the weighting function e^{-x^2}, it is possible to construct families of functions in the interval $(-\infty, \infty)$ which are orthonormal with respect to a unity weighting function. These functions are obtained by multiplying the Hermite polynomials by the square root of their weighting function, followed by a translation and a change of scale in the x axis, and, finally, by a normalization. The end result of all these operations are the so-called Hermite functions, given by

$$\mathcal{H}_n(\lambda, t_0; t) = \alpha_n H_n(x)\, e^{-x^2/2} \tag{4}$$

where $x = (t - t_0)/\lambda$ and $\alpha_n = (2^n n! \lambda \sqrt{\pi})^{-1/2}$. These functions have two free parameters: t_0, the center of the functions; and λ, their scale factor.

The Hermite functions form a complete orthonormal set in $L^2(-\infty, \infty)$, i.e., any square integrable function $f(t)$ can be expressed in the form

$$f(t) = \sum_{n=0}^{\infty} c_n(\lambda, t_0)\, \mathcal{H}_n(\lambda, t_0; t) \tag{5}$$

where

$$c_n(\lambda, t_0) = \int_{-\infty}^{\infty} f(t)\mathcal{H}_n(\lambda, t_0; t)\, dt. \tag{6}$$

The partial derivatives of the Hermite functions w.r.t. their three parameters are given by the following interesting formulas

$$\frac{\partial \mathcal{H}_n(\lambda, t_0; t)}{\partial t} = \beta_n\, \mathcal{H}_{n-1}(\lambda, t_0; t) - \beta_{n+1}\, \mathcal{H}_{n+1}(\lambda, t_0; t), \tag{7}$$

$$\frac{\partial \mathcal{H}_n(\lambda, t_0; t)}{\partial t_0} = -\frac{\partial \mathcal{H}_n(\lambda, t_0; t)}{\partial t}, \tag{8}$$

and

$$\frac{\partial \mathcal{H}_n(\lambda, t_0; t)}{\partial \lambda} = \gamma_n\, \mathcal{H}_{n-2}(\lambda, t_0; t) - \gamma_{n+2}\, \mathcal{H}_{n+2}(\lambda, t_0; t), \tag{9}$$

where

$$\beta_n = \frac{\sqrt{n}}{\lambda\sqrt{2}}, \qquad \text{and} \qquad \gamma_n = -\frac{\sqrt{n}\sqrt{n-1}}{2\lambda}.$$

These formulas express the partial derivatives of the Hermite functions as linear combinations of other Hermite functions, and can be proven easily by differentiating (4) w.r.t. the appropriate variables and simplifying the results with (1), (2), and (3). In a few words, these formulas are a consequence of the special form of the Fourier transform of the Hermite functions and of their orthonormality. Note that a similar result exists for the Laguerre functions [5], [7], [9].

The Hermite functions satisfy the following differential equation (a prime denotes differentiation w.r.t. t)

$$x^2\mathcal{H}_n(\lambda, t_0; t) - \lambda^2\mathcal{H}_n''(\lambda, t_0; t) = (2n+1)\mathcal{H}_n(\lambda, t_0; t) \tag{10}$$

[remember that $x = (t - t_0)/\lambda$]. This follows from the differentiation of (7) w.r.t. t and from (1) and (3).

3 Stationarity conditions of the squared error of a truncated Hermite series

In practice we are forced to use a finite number of terms, say $N+1$, in (5). We then obtain an approximation to $f(t)$ of the form

$$f_N(\lambda, t_0; t) = \sum_{n=0}^{N} c_n(\lambda, t_0)\, \mathcal{H}_n(\lambda, t_0; t)$$

that has a squared error of

$$\xi_N(\lambda, t_0) = \int_{-\infty}^{\infty} f^2(t)\, dt - \sum_{n=0}^{N} c_n^2(\lambda, t_0).$$

This squared error is clearly a function of the center and of the scale factor of the Hermite functions. It makes then sense to try to minimize it w.r.t. these two parameters. A preliminary step towards this goal is the deduction of the stationarity conditions of the squared error w.r.t. λ and to t_0.

Let $\theta = \lambda$ or $\theta = t_0$. Then we have

$$\frac{\partial \xi_N(\lambda, t_0)}{\partial \theta} = -2\sum_{n=0}^{N} c_n(\lambda, t_0)\frac{\partial c_n(\lambda, t_0)}{\partial \theta} \tag{11}$$

and [cf. (6)]

$$\frac{\partial c_n(\lambda, t_0)}{\partial \theta} = \int_{-\infty}^{\infty} f(t)\frac{\partial \mathcal{H}_n(\lambda, t_0; t)}{\partial \theta}\, dt.$$

Using (8) and (9) in the previous formula yields

$$\frac{\partial c_n(\lambda, t_0)}{\partial t_0} = -\beta_n\, c_{n-1}(\lambda, t_0) + \beta_{n+1}\, c_{n+1}(\lambda, t_0),$$

and

$$\frac{\partial c_n(\lambda, t_0)}{\partial \lambda} = \gamma_n\, c_{n-2}(\lambda, t_0) + \gamma_{n+2}\, c_{n+2}(\lambda, t_0).$$

Using these remarkable formulas in (11) it is possible to verify that in both cases the summation appearing there becomes a telescopic series, giving the simple results

$$\frac{\partial \xi_N(\lambda, t_0)}{\partial t_0} = -2\,\beta_{N+1}\, c_N(\lambda, t_0)\, c_{N+1}(\lambda, t_0) \tag{12}$$

and

$$\begin{aligned}\frac{\partial \xi_N(\lambda, t_0)}{\partial \lambda} &= 2\,\gamma_{N+1}\, c_{N-1}(\lambda, t_0)\, c_{N+1}(\lambda, t_0) \\ &\quad + 2\,\gamma_{N+2}\, c_N(\lambda, t_0)\, c_{N+2}(\lambda, t_0).\end{aligned} \tag{13}$$

It is easy to prove that these two partial derivatives vanish simultaneously if and only if at least one of the following three conditions is satisfied

$$c_k(\lambda, t_0) = c_{k+1}(\lambda, t_0) = 0, \quad k = N-1, N, N+1. \tag{14}$$

Intuitively, we would expect that at local minima of $\xi_N(\lambda, t_0)$ the condition corresponding to $k = N+1$ is the one satisfied, since for the other two cases one or more coefficients used in the approximation vanish. Unfortunately, this is not always the case.

If we suspect that the global minimum of $\xi_N(\lambda, t_0)$ satisfies condition (14) for a specific k, and if we have one reasonably good estimate of the optimal values of λ and t_0, then it is possible to refine these estimates using one or more iterations of Newton's method to solve a system of (two) nonlinear equations. This is possible because it is very easy to compute the partial derivatives of $\xi_N(\lambda, t_0)$ w.r.t. these two parameters [it is only necessary to compute two extra coefficients of the Hermite series, cf. (12) and (13)]. Of course, we can also use Newton's method to try to find solutions of (14) by starting at some random values of λ and t_0.

3.1 Separate approximation of the even and odd components of a function

If it is known that a function is even (or odd) around a certain point t_0 then half of the expansion coefficients

$f(t)$	m_0	m_1	m_2	m_3
$f(x)$	$m_0\,\lambda$	$m_1\,\lambda^2 + t_0 m_0\,\lambda$	$m_2\,\lambda^3 + 2t_0 m_1\,\lambda^2 + t_0^2 m_0\,\lambda$	m_3/λ

Table 1: Effect of a translation and change of scale in the measurements m_k. Note that $x = (t - t_0)/\lambda$. The second line of this table then represents a translation of $f(t)$ by t_0 followed by an amplification of the t axis by a factor of λ.

of the truncated Hermite series will be zero. In such cases it makes sense to use in the approximation only even (or odd) Hermite functions, which can be computed directly with recursion (2). Again, the partial derivative of the squared error of the approximation w.r.t. λ will be a telescopic series, and the stationarity condition will simply be

$$c_N(\lambda, t_0)\, c_{N+2}(\lambda, t_0) = 0 \qquad (15)$$

where N is the index of the last term used in the expansion.

Since even and odd functions (around the same point) are orthogonal to each other it is possible to approximate the even and odd components of a function using even and odd truncated Hermite expansions with different scale factors and with different numbers of terms (but with the same center). The two independent stationarity conditions will have the same form as (15). This idea was suggested to the authors by Dr. A. C. den Brinker.

4 Optimum value of λ and t_0 for a class of functions satisfying certain measurements

Based on the ideas of [6] (see also [10] for a more general result) it is possible to estimate a "good" center and scale factor for the Hermite functions used in the approximation of a function $f(t)$. Let

$$m_i = \int_{-\infty}^{\infty} t^i\, f^2(t)\, dt, \quad i = 0, 1, 2,$$

and

$$m_3 = \int_{-\infty}^{\infty} [f'(t)]^2\, dt = -\int_{-\infty}^{\infty} f(t) f''(t)\, dt.$$

We assume that these integrals exist and are absolutely convergent. The purpose of this section is to derive formulas for the best center and scale of the Hermite functions when the only information known about $f(t)$ is the four "measurements" $m_0, \ldots, m_3$ introduced above. Note that m_0 is the energy of $f(t)$ and that m_1 will vanish if $f(t)$ is an even or an odd function. Note also that a translation and change of scale affect these measurements in the way shown in Table 1.

Using the differential equation (10) that the Hermite functions satisfy it is possible to prove that

$$\sum_{n=0}^{\infty} (2n+1) c_n^2(\lambda, t_0) = \frac{m_2 - 2t_0\, m_1 + t_0^2\, m_0}{\lambda^2} + \lambda^2\, m_3.$$

But

$$(2N+3)\xi_N(\lambda, t_0) \le \sum_{n=N+1}^{\infty} (2n+1) c_n^2(\lambda, t_0)$$

with equality if and only if $c_n(\lambda, t_0) = 0$ for $n > N+1$. Also

$$\sum_{n=0}^{N} (2n+1) c_n^2(\lambda, t_0) \ge \sum_{n=0}^{N} c_n^2(\lambda, t_0)$$

with equality if and only if $c_n(\lambda, t_0) = 0$ for $1 \le n \le N$. Combining all these results yields

$$\xi_N(\lambda, t_0) \le \frac{m_3\, \lambda^4 - m_0\, \lambda^2 + (m_2 - 2t_0\, m_1 + t_0^2\, m_0)}{2(N+1)\lambda^2}$$

This upper bound for $\xi_N(\lambda, t_0)$ is attained if $f(t)$ is a linear combination of $\mathcal{H}_0(\lambda, t_0; t)$ and $\mathcal{H}_{N+1}(\lambda, t_0; t)$. Minimizing this bound w.r.t. t_0 yields

$$\hat{t}_0 = \frac{m_1}{m_0}, \qquad (16)$$

which is independent of λ. Hence the "best" center for the expansion can be determined quite easily and requires only two measurements. Note that (16) has a very simple and elegant interpretation, and is widely used in wavelet theory. Minimizing the upper bound also w.r.t. λ yields

$$\hat{\lambda}^4 = \frac{m_0 m_2 - m_1^2}{m_0 m_3}. \qquad (17)$$

Using the data presented in Table 1 it is reassuring to verify that translations and changes of scale of $f(t)$ are reflected in the appropriate way in the values of $\hat{t}_0$ and $\hat{\lambda}$, as we invite the reader to check. Also, for the Hermite functions themselves we have $m_0 = 1$, $m_1 = t_0$, $m_2 = \frac{2n+1}{2}\lambda^2 + t_0^2$, and $m_3 = \frac{2n+1}{2\lambda^2}$, which gives $\hat{t}_0 = t_0$ and $\hat{\lambda} = \lambda$, which is also reassuring.

These results can be extended easily to the problem of the determination of a "good" center and scale factor for the simultaneous approximation of several functions by truncated Hermite series (with the same center and scale factor). Let

$$M_k = \sum_{l=1}^{L} w_l\, m_{k,l}, \qquad k = 0, 1, 2, 3,$$

where $m_{k,l}$ is the k-th measurement for the l-th function and w_l is a positive weight that quantifies the relative importance of the l-th function. Then, estimates for a "good" center and scale factor for the Hermite functions can be obtained using (16) and (17), but replacing the (individual) lower-case m_k "measurements" by the (collective) upper-case M_k "measurements".

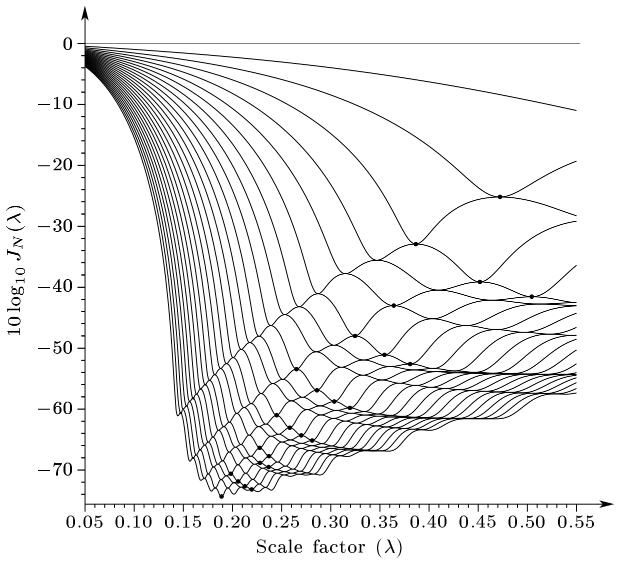

Figure 1: Normalized squared error of the approximation of $f_{1/2}(t)$ by truncated Hermite series with up to 26 even terms, i.e., for $N = 0, 2, \ldots, 50$. Note that consecutive curves touch only in stationary points (in this case, local extrema) of both curves. This is a consequence of (15).

5 An example

Consider the problem of approximating the following function (a raised cosine)

$$f_b(t) = \begin{cases} 1, & \text{if } |t| < 1-b \\ \frac{1}{2}\left[1 - \sin\left(\frac{\pi}{2}\frac{|t|-1}{b}\right)\right], & \text{if } 1-b \le |t| < 1+b \\ 0, & \text{otherwise} \end{cases}$$

by a truncated Hermite expansion. Since this is an even function around $t = 0$ we will set $t_0 = 0$ and work with Hermite series with even terms only. In Fig. 1 we depict the normalized squared error, which is given by $J_N(\lambda) = \xi_N(\lambda, 0)/m_0$, for several values of N. Although it may be dangerous to extrapolate, the figure suggests that the optimal λ approaches zero when N increases. If true, this may be explained by the fact that the Hermite functions become wider and wider when N increases, with a "width" that is roughly proportional to $\sqrt{2N+1}$ (the magnitude of their largest zero, cf. (3) and [8]). If the function being approximated has finite support, which is the case here, then it may happen that the best λ decreases to compensate for this increase in the "width" of the last Hermite functions used in the approximation.

For $f_b(t)$ it is possible to show that

$$\begin{aligned} m_0 &= \frac{4-b}{2} \\ m_1 &= 0 \\ m_2 &= \frac{(6-\pi^2)b^3 + (12\pi^2 - 96)b^2 - 3\pi^2 b + 4\pi^2}{6\pi^2} \\ m_3 &= \frac{\pi^2}{8b}, \end{aligned}$$

giving

$$\begin{aligned} \hat{t}_0 &= 0 \\ \hat{\lambda}^4 &= \frac{4b[(6-\pi^2)b^3 + (12\pi^2 - 96)b^2 - 3\pi^2 b + 4\pi^2]}{3\pi^4} \end{aligned}$$

as a "good" center and a "good" scale factor for the Hermite functions. Note that $\hat{\lambda}$ tends to zero when b tends to zero, i.e., when $f_b(t)$ tends to a rectangular function. We believe that the same phenomenon occurs for other discontinuous functions, since for these the "natural" value of m_3 is infinity. For $b = 1/2$ we have $\hat{\lambda} \approx 0.672$, which is not a good estimate of the best λ for almost all values of N (cf. Fig. 1). However, for small N, it gives a reasonably good indication of the magnitude of the optimal values of the scale factor.

References

[1] L. R. Lo Conte, R. Merletti, and G. V. Sandri. Hermite expansions of compact support waveforms: Applications to myoelectric signals. *IEEE Trans. Biomedical Engineering*, 41:1147–1159, 1994.

[2] J.-B. Martens. The Hermite transform – Theory. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 38:1595–1606, 1990.

[3] J.-B. Martens. The Hermite transform – Applications. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 38:1607–1618, 1990.

[4] G. J. Clowes. Choice of the time-scaling factor for linear system approximations using orthonormal Laguerre functions. *IEEE Trans. Automatic Control*, 10:487–489, 1965.

[5] J. J. King and T. O'Canainn. Optimum pole positions for Laguerre-function models. *Electronics Letters*, 5:601–602, 1969.

[6] T. W. Parks. Choice of time scale in Laguerre approximations using signal measurements. *IEEE Trans. Automatic Control*, 16:511–513, 1971.

[7] L. Wang and W. R. Cluett. Optimal choice of time-scaling factor for linear system approximations using Laguerre models. *IEEE Trans. Automatic Control*, 39:1463–1467, 1994.

[8] G. Szegö. *Orthogonal Polynomials*, volume XXIII of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, Rhode Island, fourth edition, 1975.

[9] T. Oliveira e Silva. On the determination of the optimal pole position of Laguerre filters. *IEEE Trans. Signal Processing*, 43:2079–2087, 1995.

[10] N. Tanguy, P. Vilbé, and L. C. Calvez. Optimum choice of free parameter in orthonormal approximations. *IEEE Trans. Automatic Control*, 40:1811–1813, 1995.