Function decomposition using distending function

József Dombi
Univeristy of Szeged
Department of Informatics
Árpád tér 2, Szeged, H–6720, Hungary
dombi@inf.u-szeged.hu

József Dániel Dombi
Univeristy of Szeged
Department of Informatics
Árpád tér 2, Szeged, H–6720, Hungary
dombijd@inf.u-szeged.hu

***Abstract*— In this paper we will present a new approach to approximate functions. The technology based on sigmoid function and Dombi operators. The main step is building a so called distending function of an interval to approximate the characteristic function. In this way we can compress the information of the function.**

I. INTRODUCTION

Function approximation (or interpolation) plays an important role. It has many applications. One of the most used algorithms is finding the regression line or more general a non-linear function with some parameters which are close to the function to be approximated. Usually ”close” means a square error. If our application is a polinom, i.e.:

$$y(\mathbf{a}, x_i) = \sum_{i=1}^{n} a_i x^i \tag{1}$$

then we want to minimize the energy function:

$$E(\mathbf{a}) = \frac{1}{2} \sum_{i=1}^{n} \left(y(x_i, \mathbf{a}) - y_i\right)^2 \tag{2}$$

If the degree of the polynom are too high (not linear) then it is not easy to interpret the optimized parameters. Our aim is to approximate the function with the help of membership like functions. We need such membership function which approximates the characteristic function. We get it by introducing the distending function modeling the inequality. Using the conjunctive operator to the distending function we get the desired function class. This result can be found in chapter 2. In chapter 3 we give two algorithms for approximating the function, and parameters are intervals and slope which one semantically interpretable.

II. DISTENDING FUNCTION

In fuzzy logic theory the membership function plays an important role. In pliant logic we use a soft inequality and we call it the distending function.

Definition 1: The distending function:

$$\delta_a^{(\lambda)}(x) = f^{-1}\left(e^{-\lambda(x-a)}\right) \qquad \lambda \epsilon R, a \epsilon R$$

Here f is the generator function of the logical connectives, λ is responsible for the sharpness and a is the threshold value.

Definition 2: The pliant system is strict monotonisly incresing t-norm and t-conorm. The following expression is valid for the generator function:

$$f_c(x) \cdot f_d(x) = 1$$

The semantic meaning of $\delta_a^{(\lambda)}$ is

$$truth(a <_\lambda x) = \delta_a^{(\lambda)}(x)$$

Remark 1:

1) In pliant system f could be the generator function of the conjunctive operator or the disjunctive operator. The form of $\delta_a^{(\lambda)}(x)$ is the same in both cases.
2) In the pliant concept the operators and membership are closely related.
3) Using the soft inequality with the distending function cannot describe a membership like ”middle age”.

The distending function in the Dombi operator case is the sigmoid (logistic) function:

$$\sigma_a^{(\lambda)}(x) = \frac{1}{1 + e^{-\lambda(x-a)}}$$

A. Distending interval

In fuzzy concept membership function has different interpretation. In pliant concept membership function replaced by soft interval. It's mathematical description is

$$\delta_{a,b}^{\lambda_1,\lambda_2}(x) = truth(a <_{\lambda_1} x <_{\lambda_2} b)$$

Using the pliant concept we translate it into two inequalities corresponding it by an ”and” (conjunctive) operator.

$$truth(a <_{\lambda_1} x) \; and \; truth(x <_{\lambda_2} b)$$

Theorem 1: In pliant system the distending interval if

$$\delta_{a,b}^{\lambda_1,\lambda_2}(a) = \nu_0 \qquad \delta_{a,b}^{\lambda_1,\lambda_2}(b) = \nu_0 \tag{3}$$

is

$$\delta_{a,b}^{\lambda_1,\lambda_2}(x) = f^{-1}\left(\frac{1}{A}\left(A_1 e^{-\lambda_1(x-a)} + A_2 e^{-\lambda_2(b-x)}\right)\right) \tag{4}$$

where

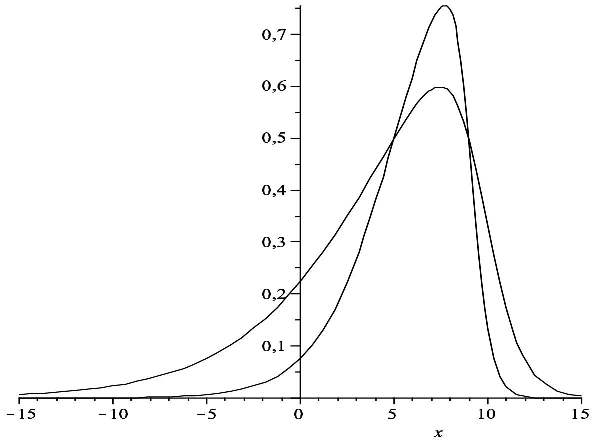

Fig. 1. $\sigma_{a,b}^{\lambda_1,\lambda_2}(x)$ if $a = 5, b = 9, \nu_0 = \frac{1}{2}, \lambda_1 = \frac{1}{2}, \lambda_2 = 2$ and $\lambda_1 = \frac{1}{4}, \lambda_2 = 1$

$$\begin{aligned} A &= \frac{1}{f(\nu_0)}\left(1 - e^{-(\lambda_1+\lambda_2)(b-a)}\right) \\ A_1 &= 1 - e^{-\lambda_2(b-a)} \\ A_2 &= 1 - e^{-\lambda_1(b-a)} \end{aligned} \tag{5}$$

Proof 1: It is a strike forward calculation.
At the Dombi operator case:

$$\sigma_{a,b}^{\lambda_1,\lambda_2}(x) = \frac{1}{1 + \frac{1-\nu_0}{\nu_0}\frac{1}{A}\left(A_1 e^{-\lambda_1(x-a)} + A_2 e^{-\lambda_2(b-x)}\right)} \tag{6}$$

where

$$\begin{aligned} A &= 1 - e^{-(\lambda_1+\lambda_2)(b-a)} \\ A_1 &= 1 - e^{-\lambda_2(b-a)} \\ A_2 &= 1 - e^{-\lambda_1(b-a)} \end{aligned}$$

In Figure (1) we have plotted $\sigma_{a,b}^{\lambda_1,\lambda_2}(x)$ using different parameter values.

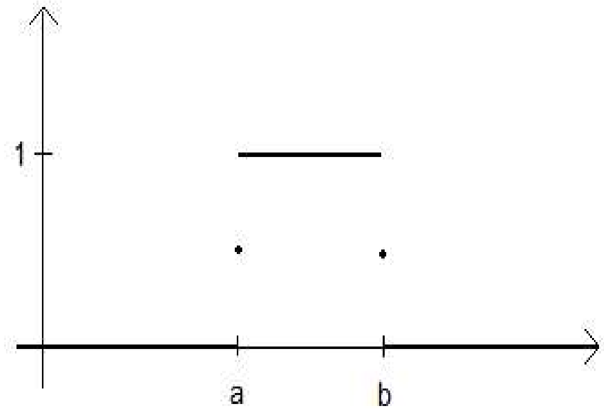

Fig. 2. $\delta_{a,b}(x)$ function

The following properties hold for the distending interval:

Theorem 2:

$$\delta_{a,b}(x) = \lim_{\lambda_1\to\infty,\lambda_2\to\infty} \delta_{a,b}^{\lambda_1,\lambda_2}(x) = \begin{cases} 0 & \text{if} \quad x < a \\ \nu_0 & \text{if} \quad x = a \\ 1 & \text{if} \quad a < x < b \\ \nu_0 & \text{if} \quad x = b \\ 0 & \text{if} \quad b < x \end{cases} \tag{7}$$

See Figure (2)

Proof 2: Because $\delta_{a,b}^{\lambda_1,\lambda_2}(a) = \delta_{a,b}^{\lambda_1,\lambda_2}(b) = \nu_0$ we have to prove just 3 cases. Here we will prove the $a < x < b$ case. The other two cases can be proved in a similar way.
It is obvious because the properties of the exponential function.

From (6) we can derive another type of distending function, where $a = b$. This means that not the thresholds of value a and b are given although the value a where the maximum is taken.

Theorem 3: The following limes property holds:

$$\delta_{a}^{\lambda_1,\lambda_2}(x) = \lim_{a\to b} \delta_{a,b}^{\lambda_1,\lambda_2}(x) = \tag{8}$$

$$f^{-1}\left(f(\nu_0)\left(\frac{\lambda_2}{\lambda_1+\lambda_2}e^{-\lambda_1(x-a)} + \frac{\lambda_1}{\lambda_1+\lambda_2}e^{-\lambda_2(a-x)}\right)\right)$$

At the Dombi operator case

$$\sigma_{a}^{\lambda_1,\lambda_2}(x) = \lim_{a\to b} \sigma_{a,b}^{\lambda_1,\lambda_2}(x) = \tag{9}$$

$$\frac{1}{1 + \frac{1-\nu_0}{\nu_0}\left(\frac{\lambda_2}{\lambda_1+\lambda_2}e^{-\lambda_1(x-a)} + \frac{\lambda_1}{\lambda_1+\lambda_2}e^{-\lambda_2(a-x)}\right)}$$

In Figure (3) $\sigma_{a}^{\lambda_1,\lambda_2}(x)$ is shown with typical values.

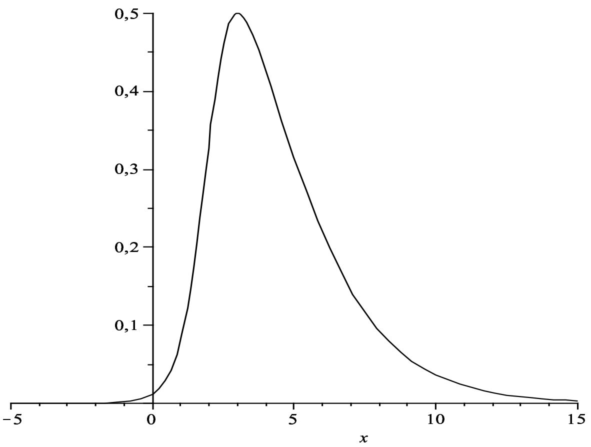

Fig. 3. $\sigma_a^{\lambda_1,\lambda_2}(x)$ if $a = 3, \nu_0 = \frac{1}{2}, \lambda_1 = 2, \lambda_2 = \frac{1}{2}$

We can get an impulse function from $\delta_a^{\lambda_1,\lambda_2}(x)$.

Theorem 4:

$$\delta_a(x) = \lim_{\lambda_1,\lambda_2\to\infty} \delta_a^{\lambda_1,\lambda_2}(x) = \begin{cases} \nu_0 & \text{if} \quad x = a \\ 0 & \text{if} \quad x \neq a \end{cases} \tag{10}$$

Now we can use Equation (6) if a and b are given and Equation (8) if the maximum point a is given.

III. Construction of the approximation

If the functions belong to the integrable function in the Riemannian sense, then there exist upper or lower approximations of rectangles.

We will use this fact in the next theorem.

Theorem 5: Let $f(x)$ integrable function in the Riemannishe sense, and let $a_1 < a_2 \ldots < a_n$ discretization of the interval of the domain of the approximated function and let

$$G(x) = \sum_{i=1}^{m} y_i \delta_{a_i,a_{i+1}}^{\lambda_i,\lambda_{i+1}}(x) \tag{11}$$

then $\int \|f(x) - G(x)\| \rightarrow 0$ if $max\,\|a_{i+1} - a_i\| \rightarrow 0$ and $\lambda_i \rightarrow \infty$.

See Figure (4).

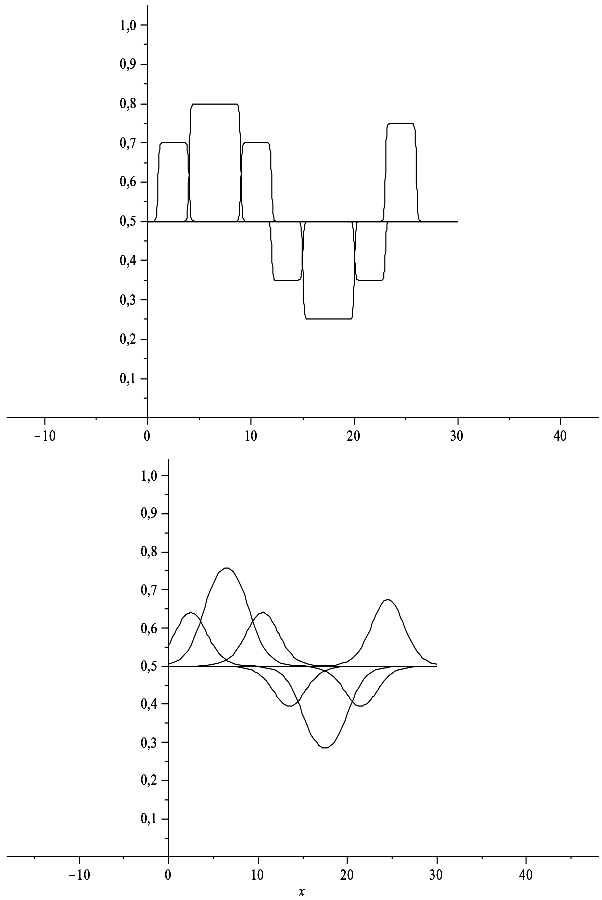

Fig. 4. Rectangles without aggregation, $\lambda = 16$ and $\lambda = 1$

We can use the impulse function to interpolate the function: If λ_1, λ_2 are not too large then we get a smooth approximation. Similar theorem 5 is valid if we use impulse function

$$H(x) = \sum_{i=1}^{m} y_i \delta_{a_i}^{\lambda_1,\lambda_2}(x) \tag{12}$$

In Figure (5) we show the rectangle approximation of a function, if $\lambda = 16$ and $\lambda = 1$, while in Figure (7) we show the interpolation when $\lambda = 16$ and $\lambda = 1$.

IV. Function decomposition

In the previous section we saw that we can construct and approximate a desired function using Equation 11 and 12. We will show that how we can do the approximation by using a global optimization method. If the initial values are chosen well, it is enough to call a local search algorithm. We will use the well-known BFGS method [9].

At the practical application we have coordinates. In this work we use function with a dense sampling procedure. In all example we have using 100 equidistant coordinates on the given interval.

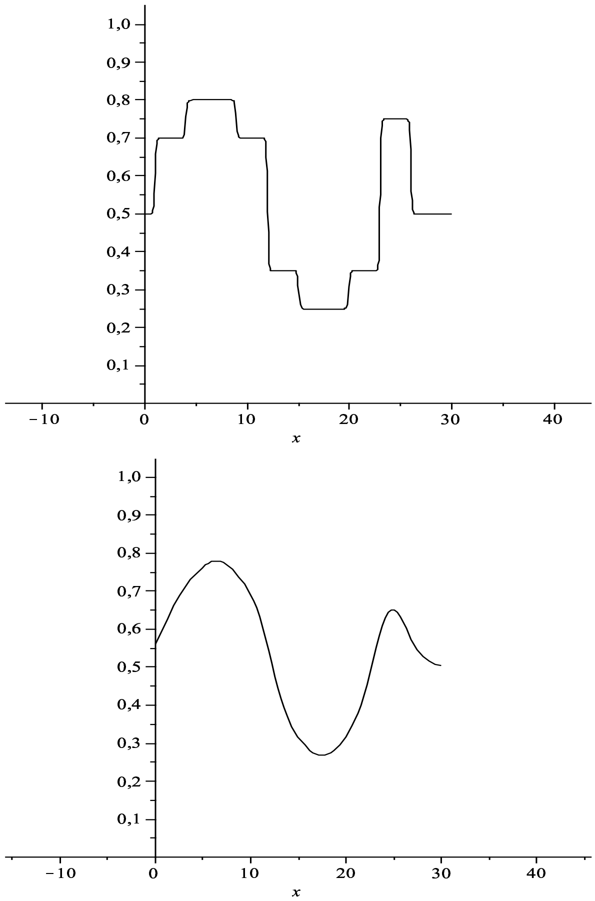

Fig. 5. Rectangles with aggregation if $\lambda = 16$ and $\lambda = 1$

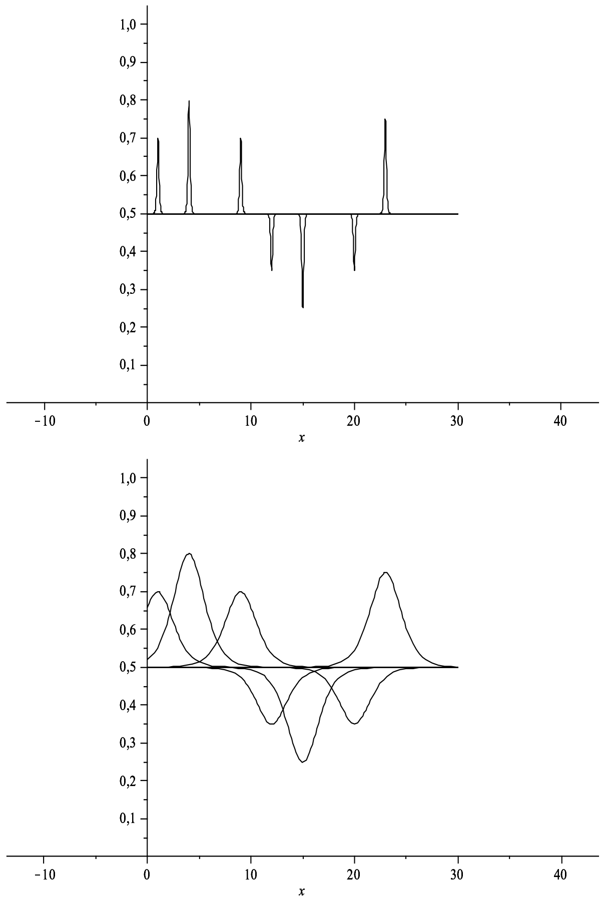

Fig. 6. Interpolative approximation without aggregation if $\lambda = 16$ and $\lambda = 1$

The global procedure tries to find all the effects simultaneously.

Let given a function $F : R \to [0, 1]$ to be approximate. Our task is to decompose it into effects. It can be done via our approximation method or interpolation procedures. First the common step is to smooth the function $F(x)$.

Algorithm for the interval approximation:

1) Let us find the the local minimum and maximum value of $F(x)$ function

$$F(c_i) = A_i \qquad \text{so that} \quad F(x) < A_i \quad \text{if} \quad x \epsilon (c_i - \varepsilon, c_i + \varepsilon)$$

$$F(c_j) = A_j \qquad \text{so that} \quad F(x) > A_j \quad \text{if} \quad x \epsilon (c_i - \varepsilon, c_i + \varepsilon)$$

2) Let us define the $[a_i, b_i]$ intervals

$$a_1 = c_1 - \frac{c_1 + c_2}{2}, b_1 = \frac{c_1 + c_2}{2},$$
$$a_2 = \frac{c_1 + c_2}{2}, b_2 = \frac{c_2 + c_3}{2} \dots$$
$$a_n = \frac{c_{n-1} + c_n}{2}, b_n = c_n + \frac{c_{n-1} + c_n}{2}$$

where

$$c_1 < c_2 < c_3 < \dots < c_k$$

We suppose that at the maximum (minimum) value appears, like as shown in Figure (8) below.

3) Let us define the initial value of λ_{i_1} and λ_{i_2}

$$\lambda_{i_1} = \frac{f(c_i) - f(a_i)}{c_i - a_i} \qquad \lambda_{i_2} = 2\frac{f(b_i) - f(c_i)}{b_i - a_i}$$

4) Let us build the initial values of the function and use 11 and 12

$$E_i(x) = E_{a_i, b_i}^{\lambda_{i_1}, \lambda_{i_2}}(\gamma, x)$$

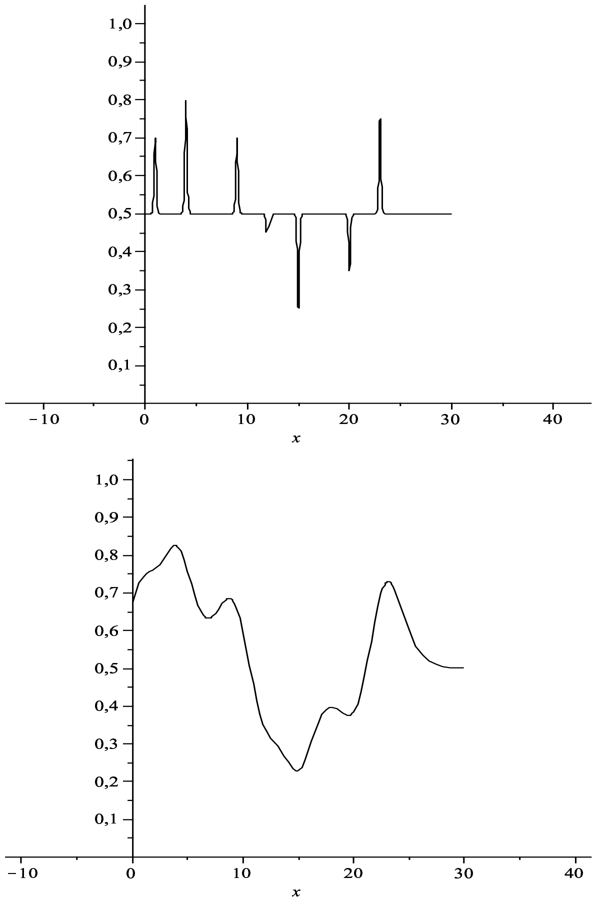

Fig. 7. Aggregation of positive and negative effects (interpolation) if $\lambda = 16$ and $\lambda = 1$

5) Find the optimal solution of the $a_i, b_i, \underline{\gamma}, \lambda_{i_1}, \lambda_{i_2}$ values with the suggested initial values.

$$\min_{\mathbf{a},\mathbf{b},\underline{\lambda}_1,\underline{\lambda}_2} \sum \left(G_{\mathbf{a},\mathbf{b}}^{\underline{\lambda}_1,\underline{\lambda}_2}(x_i) - F(x_i) \right)^2$$

It is not easy to minimize because a minimum may not be the global minimum. However, because $G_{\mathbf{a},\mathbf{b}}^{\underline{\lambda}_1,\underline{\lambda}_2}(x)$ is a continuous function of its parameters and the initial values are well-chosen, we can get good results.

The results of this are presented in Figure (10).

Algorithm for the impule approximation:

Let us find the maximum and minimum values of $F(x)$

$$c_1 < c_2 < c_3 < \ldots < c_k$$

where c_i, c_{i+1} minimum, maximum (or maximum, minimum) points. (See Figure (11)).

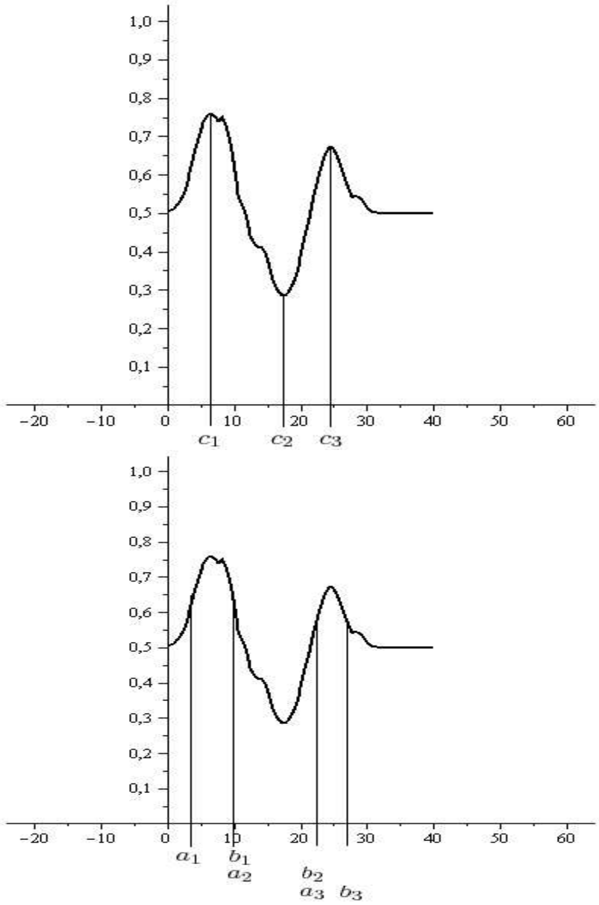

Fig. 8. Extreme values and its intervals

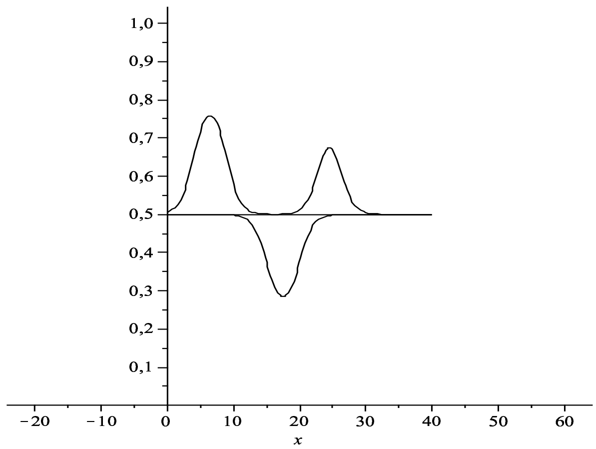

Fig. 9. Optimal components

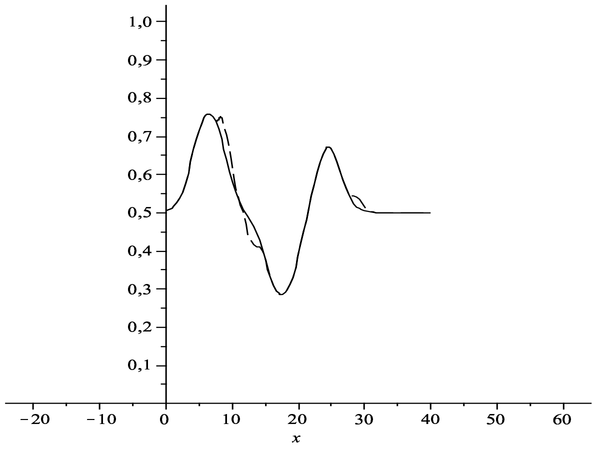

Fig. 10. The function and its approximation

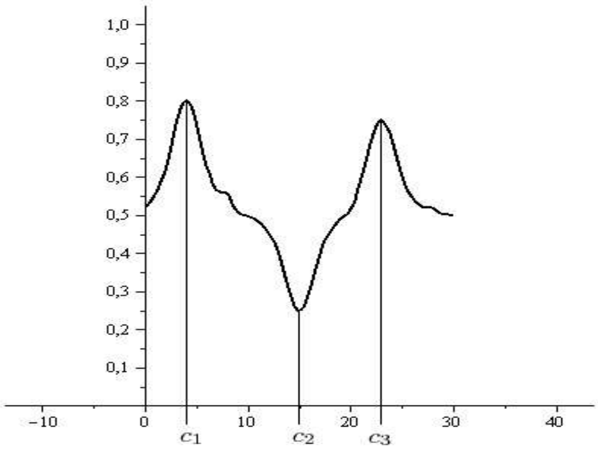

Fig. 11. Extreme values of the function

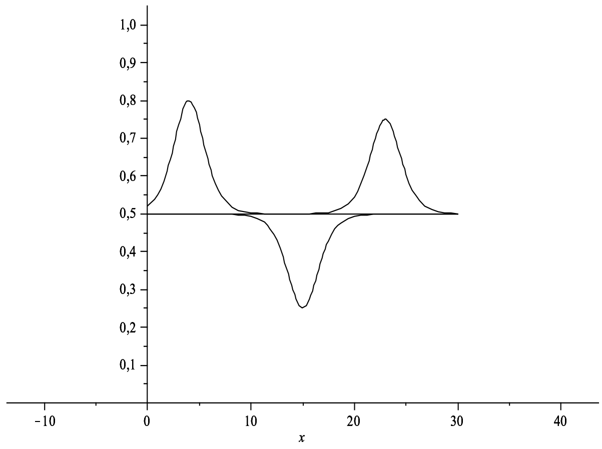

Fig. 12. Main and optimal effects for the interpolative case

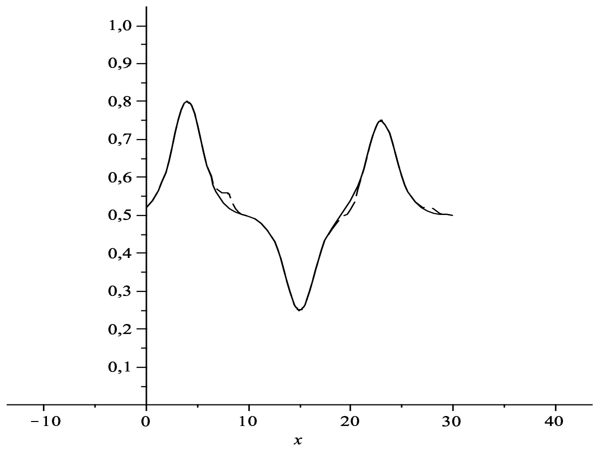

Fig. 13. The function and its interpolative approximation

And $f(c_i) = A_i$ then let the initial value of the approximation function the following

$$A_i - \frac{1}{2}, \qquad \lambda_{1_i} = \frac{c_i - c_{i-1}}{A_i - A_{i-1}} \qquad \text{and} \qquad \lambda_{2_i} = \frac{c_{i+1} - c_i}{A_{i+1} - A_i}$$

The procedure here is the same as that for the interval approximation.

In Figure (13) plotted we can see the result of applying this procedure.

REFERENCES

[1] J. Dombi: A general class of fuzzy operators, the De Morgan class of fuzzy operators and fuzziness measures induced by fuzzy operators, Fuzzy Sets and Systems Vol. 8., 149-163, 1982

[2] J. Dombi: Basic concept for a theory of evaluation: the aggregation operator, Europian Journal of Operations Research Vol. 10., 282–293, 1982

[3] M. Detyniecki, Mathematical aggregation operators and their application to video querying, Doctoral thesis - research report 2001-002, Laboratoire d'Informatique de Paris 6, 2000

[4] J. Dombi: The pliant concept and the Generalized Dombi operator. In: Proceedings of IEEE International Workshop on Soft Computing Applications, Szeged-Hungary and Arad-Romania (2005) 93-104.

[5] J. Fodor, R.R. Yager, and A. Rybalov. Structure of uninorms. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems Vol. 5., 411-427, 1997

[6] D. Dubois, H. Prade: Fundamentals of fuzzy sets (The hanbooks of fuzzy sets series), Kluwer, 2000

[7] R.R. Yager. Uninorms in fuzzy systems modeling. Fuzzy Sets and Systems, Vol. 122., 167-175, 2001

[8] Andrzej Ruszczynski: Nonlinear Optimization, Princeton University Press, 2006

[9] J. Dombi: Towards a General Class of Operators for Fuzzy Systems. IEEE Transaction on Fuzzy Systems, Vol. 16., 477-484, 2008

[10] J. Aczl: Lectures on Functional Equations and Applications, Academic Press, New York, 1966.