Available online at www.sciencedirect.com

SCIENCE DIRECT®

Fuzzy Sets and Systems 154 (2005) 275–286

FUZZY sets and systems

www.elsevier.com/locate/fss

The approximation of piecewise linear membership functions and Łukasiewicz operators

József Dombi, Zsolt Gera*

University of Szeged, Institute of Informatics, 6701 Pf. 652, Szeged, Hungary

Received 18 November 2004; received in revised form 22 February 2005; accepted 22 February 2005
Available online 11 March 2005

Abstract

In this paper we propose an approximation of piecewise linear membership functions with the help of sigmoid functions and certain arithmetic operations. The gradient-based tuning of piecewise linear membership functions can be achieved with the proposed efficient approximation because it has simple continuous derivatives. With this construction we can even approximate the Łukasiewicz operator family which plays an important role in fuzzy logic, first of all from the theoretical point of view, although in practice in optimization and learning it is rarely used because the lack of good analytical properties, e.g. a continuous gradient. The proposed approximation enlarges the applicability of fuzzy methods to the operators and membership functions where the differentiability is desirable.

Keywords: Fuzzy connectives; Membership function; Approximation; Łukasiewicz operators; Nilpotent operators

1. Introduction

The construction and the interpretation of fuzzy membership functions have always been a crucial question. Bilgic and Türksen gave a comprehensive overview of the most relevant interpretations in [11]. For the construction of membership functions Dombi [10] had an axiomatic point of view, Civanlar and Trussel [8] used statistical data, Bagis [2], Denna et al. [9], Karaboga [14] applied tabu search. However, most fuzzy applications use piecewise linear membership functions because of their easy handling, for example in embedded fuzzy control applications where the limited computational resources does not

* Corresponding author. Tel.: +36 62544222/3830.
E-mail address: gera@inf.u-szeged.hu (Z. Gera).

0165-0114/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.fss.2005.02.016

allow the use of complicated membership functions. In other areas where the model parameters are learned by a gradient-based optimization method, they cannot be used because the lack of continuous derivatives. For example, to fine tune a fuzzy control system by a simple gradient-based technique it is required that the membership functions are differentiable for every input. There are numerous papers dealing with the concept of fuzzy set approximation and membership function differentiability (see for e.g. [3,12,16]). In this paper we give a different solution to the problem of non-differentiability of piecewise linear functions by approximating the "cut" function of the Łukasiewicz operators, and use it to construct continuously differentiable membership functions which approach the well-known triangular or trapezoidal membership functions.

The paper is organized as follows: Section 2 is a brief overview of Łukasiewicz operators, in Section 3 we give the basic properties of the sigmoid function which serves as the basis for the approximation, prove the main approximation theorem, and examine the derivatives and the convergence of the proposed approximation and in Section 4 we apply the approximation to triangular and trapezoidal membership functions.

2. Łukasiewicz operators

The Łukasiewicz operator class (see e.g. [1,7,13]) is commonly used for various purposes, see e.g. [4,5]. In this well-known operator family the cut function (denoted by $[\cdot]$) plays an important role. We can get the cut function from x by taking the maximum of 0 and x and then taking the minimum of the result and 1.

Definition 1. Let the cut function be

$$[x] = \min(\max(0, x), 1) = \begin{cases} 0 & \text{if } x \leqslant 0, \\ x & \text{if } 0 < x < 1, \\ 1 & \text{if } 1 \leqslant x. \end{cases}$$

Let the generalized cut function be

$$[x]_{a,b} = [(x-a)/(b-a)] = \begin{cases} 0 & \text{if } x \leqslant a, \\ \dfrac{x-a}{b-a} & \text{if } a < x < b, \\ 1 & \text{if } b \leqslant x, \end{cases}$$

where $a, b \in \mathbb{R}$ and $a < b$.

In neural networks terminology this cut function is called saturating linear transfer function. All nilpotent operators are constructed using the cut function. The formulas of the nilpotent conjunction, disjunction, implication and negation are the following:

$$\begin{aligned} c(x, y) &= [x + y - 1], \\ d(x, y) &= [x + y], \\ i(x, y) &= [1 - x + y], \\ n(x) &= 1 - x, \end{aligned} \tag{1}$$

where $x, y \in [0, 1]$. The truth tables of the former three can be seen in Fig. 1.

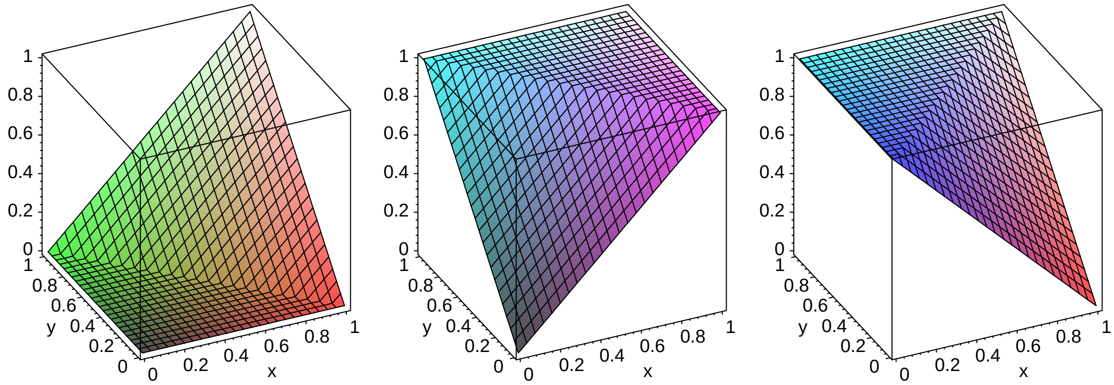

Fig. 1. The truth tables of the nilpotent conjunction, disjunction and implication.

Throughout this paper we will refer to triangular and trapezoidal membership functions as piecewise linear membership functions. They are very common in fuzzy control because of their easy handling. The generalized cut function can be used to describe piecewise linear membership functions. Generally, a trapezoidal membership function can be constructed as the conjunction of two generalized cut functions as

$$\begin{aligned} c([x]_{a,b}, 1-[x]_{c,d}) &= [[x]_{a,b} + 1 - [x]_{c,d} - 1] \\ &= [[x]_{a,b} - [x]_{c,d}], \end{aligned} \tag{2}$$

where a, b, c, d are real numbers and $a < b \leqslant c < d$. As a special case, if $b = c$ then we get a triangular membership function. For an example of the general case see Fig. 2.

The Łukasiewicz operator family used above has good theoretical properties. These are, for example, the law of non-contradiction (that is the conjunction of a variable and its negation is always zero) and the law of excluded middle (that is the disjunction of a variable and its negation is always one) both hold, and the residual and material implications coincide. These properties make these operators to be widely used in fuzzy logic and to be the closest one to classic Boolean logic. Besides these good theoretical properties this operator family does not have a continuous gradient. So for example gradient-based optimization techniques are impossible with Łukasiewicz operators. The root of this problem is the shape of the cut function itself.

Currently, we are working on new concept called pliant system in which we build up our theoretical framework by the practical (application oriented) point of view. We are using a new notation system too. Some elements of it appear in this work as well.

3. Approximation of the cut function

A solution to the above-mentioned problem is a continuously differentiable approximation of the cut function, which can be seen in Fig. 3. In this section we will construct such an approximating function by means of sigmoid functions. The reason for choosing the sigmoid function was that this function has

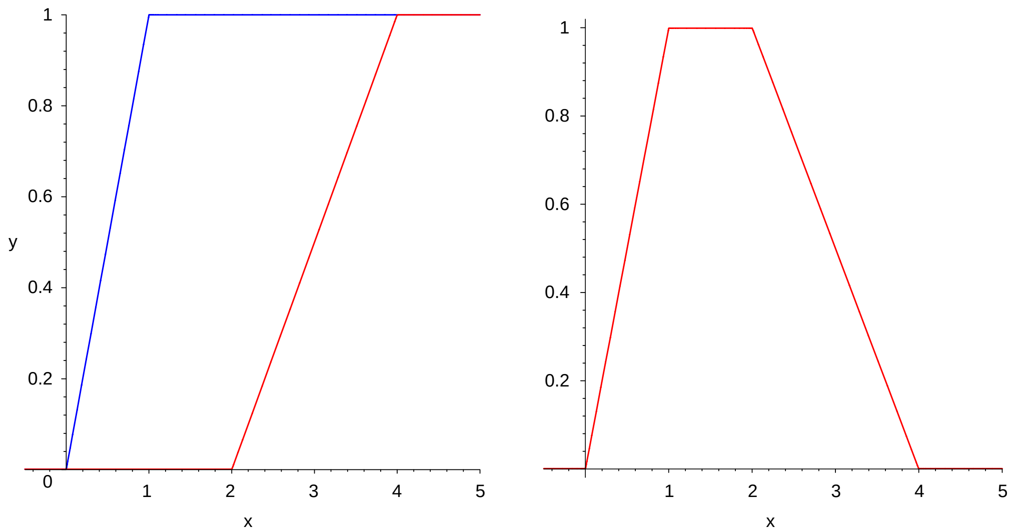

Fig. 2. In the first image: two generalized cut functions. In the second image: a trapezoidal membership function constructed as the conjunction of the former two, with a negation applied to the right one. The parameters were the following: $a = 0$, $b = 1$, $c = 2$, $d = 4$.

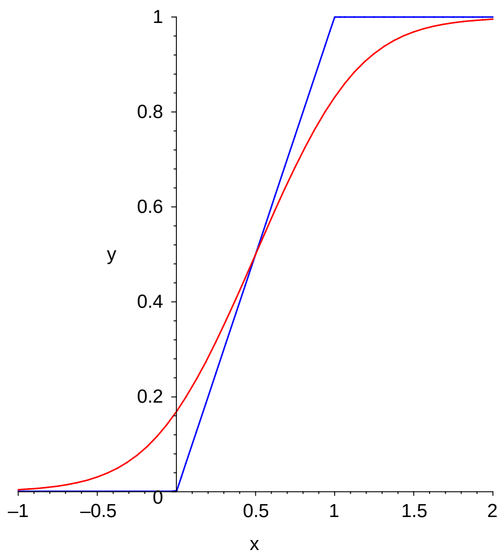

Fig. 3. The cut function and its approximation.

a very important role in many areas. It is frequently used in artificial neural networks [6], optimization methods, economical and biological models [15].

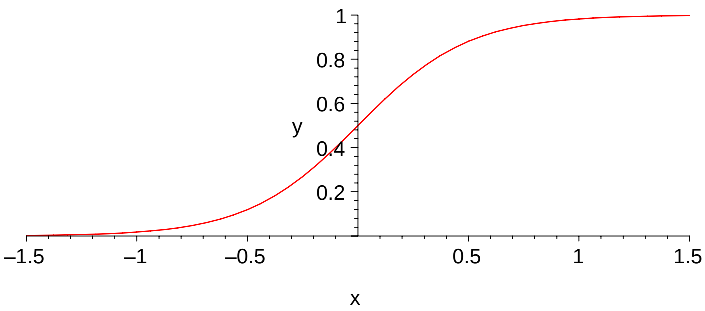

Fig. 4. The sigmoid function, with parameters $d = 0$ and $\beta = 4$.

3.1. *The sigmoid function*

The sigmoid function (see Fig. 4) is defined as

$$\sigma_d^{(\beta)}(x) = \frac{1}{1 + \mathrm{e}^{-\beta(x-d)}}, \tag{3}$$

where the lower index d is omitted if 0.

Let us examine some of its properties which will be useful later:

- Its derivative can be expressed by itself (see Fig. 5):

$$\frac{\partial \sigma_d^{(\beta)}(x)}{\partial x} = \beta \sigma_d^{(\beta)}(x) \left(1 - \sigma_d^{(\beta)}(x)\right). \tag{4}$$

- Its integral has the following form:

$$\int \sigma_d^{(\beta)}(x)\, \mathrm{d}x = -\frac{1}{\beta} \ln\left(\sigma_d^{(-\beta)}(x)\right). \tag{5}$$

Because the sigmoid function is asymptotically 1 as x tends to infinity, the integral of the sigmoid function is asymptotically x (see Fig. 6).

3.2. *The interval $[a, b]$ squashing function*

In order to get an approximation of the generalized cut function, let us integrate the difference of two sigmoid functions, which are translated by a and b ($a < b$), respectively.

$$\begin{aligned} &\frac{1}{b-a} \int \left(\sigma_a^{(\beta)}(x) - \sigma_b^{(\beta)}(x)\right) \mathrm{d}x \\ &\quad = \frac{1}{b-a} \left(\int \sigma_a^{(\beta)}(x)\, \mathrm{d}x - \int \sigma_b^{(\beta)}(x)\, \mathrm{d}x\right) \\ &\quad = \frac{1}{b-a} \left(-\frac{1}{\beta} \ln\left(\sigma_a^{(-\beta)}(x)\right) + \frac{1}{\beta} \ln\left(\sigma_b^{(-\beta)}(x)\right)\right). \end{aligned} \tag{6}$$

After simplification we get the interval $[a, b]$ squashing function:

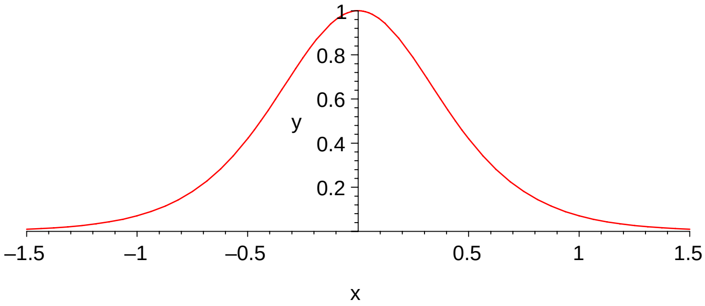

Fig. 5. The first derivative of the sigmoid function.

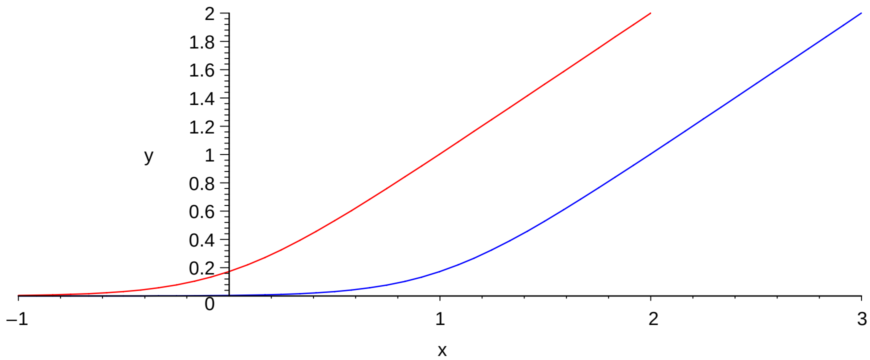

Fig. 6. The integral of the sigmoid function, one is shifted by 1.

Definition 2. Let the interval $[a, b]$ squashing function be

$$S_{a,b}^{(\beta)}(x) = \frac{1}{b-a} \ln\left(\frac{\sigma_b^{(-\beta)}(x)}{\sigma_a^{(-\beta)}(x)}\right)^{1/\beta} = \frac{1}{b-a} \ln\left(\frac{1+\mathrm{e}^{\beta(x-a)}}{1+\mathrm{e}^{\beta(x-b)}}\right)^{1/\beta}.$$

The parameters a and b affect the placement of the interval squashing function, while the β parameter drives the precision of the approximation. We need to prove that $S_{a,b}^{(\beta)}(x)$ is really an approximation of the generalized cut function.

Theorem 3. *Let $a, b \in \mathbb{R}$, $a < b$ and $\beta \in \mathbb{R}^+$. Then*

$$\lim_{\beta\to\infty} S_{a,b}^{(\beta)}(x) = [x]_{a,b}$$

and $S_{a,b}^{(\beta)}(x)$ is continuous in x, a, b and β.

Proof. It is easy to see the continuity, because $S_{a,b}^{(\beta)}(x)$ is a simple composition of continuous functions and because the sigmoid function has a range of [0, 1] the quotient is always positive.

In proving the limit we separate three cases, depending on the relation between a, b and x.

Case 1: ($x < a < b$): Since $\beta(x-a) < 0$, so $\mathrm{e}^{\beta(x-a)} \to 0$ and similarly $\mathrm{e}^{\beta(x-b)} \to 0$. Hence the quotient converges to 1 if $\beta \to \infty$, and the logarithm of one is zero.

Case 2: ($a \leqslant x \leqslant b$):

$$\begin{aligned}
&\frac{1}{b-a}\ \ln\left(\lim_{\beta\to\infty}\left(\frac{1+\mathrm{e}^{\beta(x-a)}}{1+\mathrm{e}^{\beta(x-b)}}\right)^{1/\beta}\right)\\
&\quad=\frac{1}{b-a}\ \ln\left(\lim_{\beta\to\infty}\left(\frac{\mathrm{e}^{\beta(x-a)}(\mathrm{e}^{-\beta(x-a)}+1)}{(1+\mathrm{e}^{\beta(x-b)})}\right)^{1/\beta}\right)\\
&\quad=\frac{1}{b-a}\ \ln\left(\lim_{\beta\to\infty}\frac{\mathrm{e}^{x-a}(\mathrm{e}^{-\beta(x-a)}+1)^{1/\beta}}{(1+\mathrm{e}^{\beta(x-b)})^{1/\beta}}\right)\\
&\quad=\frac{1}{b-a}\ \ln\left(\mathrm{e}^{x-a}\lim_{\beta\to\infty}\frac{(\mathrm{e}^{-\beta(x-a)}+1)^{1/\beta}}{(1+\mathrm{e}^{\beta(x-b)})^{1/\beta}}\right).
\end{aligned}$$

We transform the nominator so that we can take the e^{x-a} out of the limes. In the nominator $\mathrm{e}^{-\beta(x-a)}$ remained which converges to 0 as well as $\mathrm{e}^{\beta(x-b)}$ in the denominator so the quotient converges to 1 if $\beta \to \infty$. So as the result, the limit of the interval squashing function is $(x-a)/(b-a)$, which by definition equals to the generalized cut function in this case.

Case 3: ($a < b < x$):

$$\begin{aligned}
&\frac{1}{b-a}\ \ln\left(\lim_{\beta\to\infty}\left(\frac{1+\mathrm{e}^{\beta(x-a)}}{1+\mathrm{e}^{\beta(x-b)}}\right)^{1/\beta}\right)\\
&\quad=\frac{1}{b-a}\ \ln\left(\lim_{\beta\to\infty}\left(\frac{\mathrm{e}^{\beta(x-a)}(\mathrm{e}^{-\beta(x-a)}+1)}{\mathrm{e}^{\beta(x-b)}(\mathrm{e}^{-\beta(x-b)}+1)}\right)^{1/\beta}\right)\\
&\quad=\frac{1}{b-a}\ \ln\left(\lim_{\beta\to\infty}\frac{\mathrm{e}^{x-a}(\mathrm{e}^{-\beta(x-a)}+1)^{1/\beta}}{\mathrm{e}^{x-b}(\mathrm{e}^{-\beta(x-b)}+1)^{1/\beta}}\right)\\
&\quad=\frac{1}{b-a}\ \ln\left(\frac{\mathrm{e}^{x-a}}{\mathrm{e}^{x-b}}\lim_{\beta\to\infty}\frac{(\mathrm{e}^{-\beta(x-a)}+1)^{1/\beta}}{(\mathrm{e}^{-\beta(x-b)}+1)^{1/\beta}}\right).
\end{aligned}$$

We do the same transformations as in the previous case but we take e^{x-b} from the denominator, too. After these transformations the remaining quotient converges to 1, so

$$\begin{aligned}
\lim_{\beta\to\infty} S_{a,b}^{(\beta)}(x) &= \frac{1}{b-a}\ \ln\left(\frac{\mathrm{e}^{x-a}}{\mathrm{e}^{x-b}}\right)\\
&= \frac{1}{b-a}\ \ln(\mathrm{e}^{x-a-(x-b)})
\end{aligned}$$

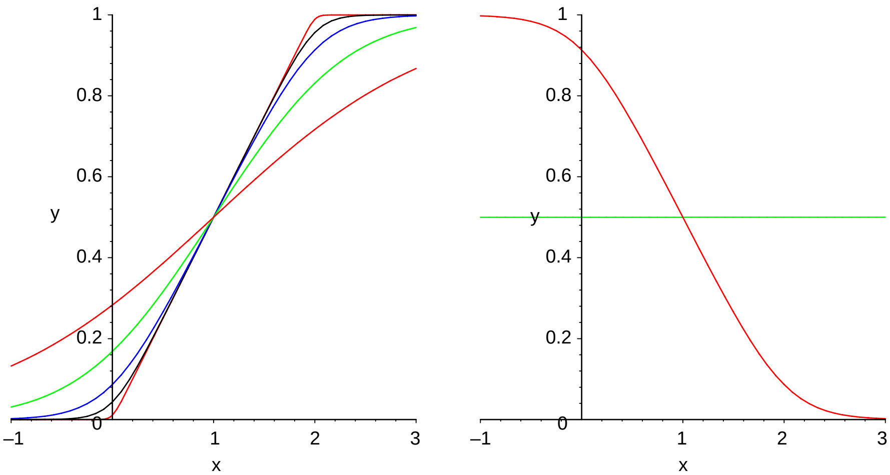

Fig. 7. In the first image: the interval squashing function with an increasing β parameter ($a = 0$ and $b = 2$). In the second image: the interval squashing function with a zero and a negative β parameter.

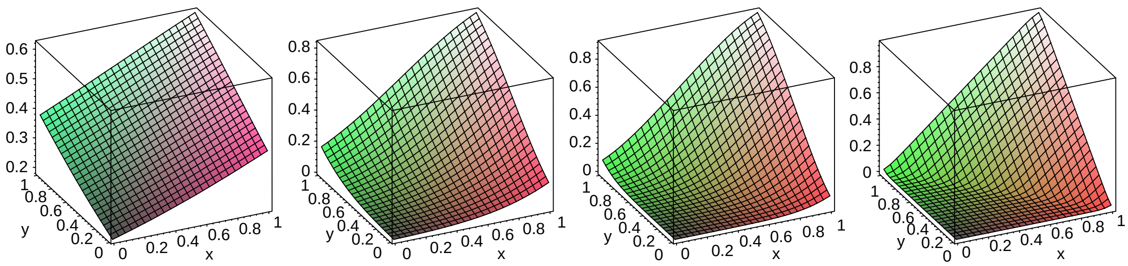

Fig. 8. The approximation of the nilpotent conjunction with β values 1, 4, 8 and 32.

$$= \frac{1}{b-a} \ln(e^{b-a}) = \frac{b-a}{b-a} = 1. \qquad \square$$

In Fig. 7 the interval squashing function can be seen with various β parameters. The following proposition states some properties of the interval squashing function.

Proposition 4.

$$\lim_{\beta \to 0} S_{a,b}^{(\beta)}(x) = 1/2,$$

$$S_{a,b}^{(-\beta)}(x) = 1 - S_{a,b}^{(\beta)}.$$

As an another example, the nilpotent conjunction is approximated with the interval squashing function in Fig. 8.

For further use, let us introduce an another form of the interval squashing function's formula. Instead of using parameters a and b which were the "bounds" on the x-axis, from now on we will use a and δ,

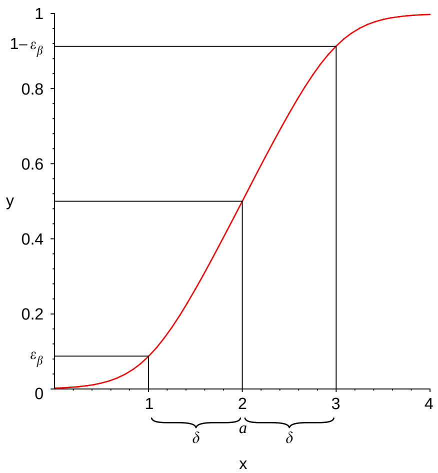

Fig. 9. The meaning of $\langle a <_\delta x\rangle_\beta$.

where a gives the center of the squashing function and where δ gives its steepness. Together with the new formula we introduce its pliant notation.

Definition 5. Let the squashing function be

$$\langle a <_\delta x\rangle_\beta = S_{a,\delta}^{(\beta)}(x) = \frac{1}{2\delta} \ln \left(\frac{\sigma_{a+\delta}^{(-\beta)}(x)}{\sigma_{a-\delta}^{(-\beta)}(x)} \right)^{1/\beta},$$

where $a \in \mathbb{R}$ and $\delta \in \mathbb{R}^+$.

If the a and δ parameters are both $1/2$ we will use the following pliant notation for simplicity:

$$\langle x\rangle_\beta = S_{1/2,1/2}^{(\beta)}(x),$$

which is the approximation of the cut function.

The inequality relation in the pliant notation refers to the fact that the squashing function can be interpreted as the truthness of the relation $a < x$ with decision level $1/2$, according to a fuzziness parameter δ and an approximation parameter β (see Fig. 9).

The derivatives of the squashing function can be expressed by itself and sigmoid functions:

$$\frac{\partial S_{a,\delta}^{(\beta)}(x)}{\partial x} = \frac{1}{2\delta} \left(\sigma_{a-\delta}^{(\beta)}(x) - \sigma_{a+\delta}^{(\beta)}(x) \right), \tag{7}$$

$$\frac{\partial S_{a,\delta}^{(\beta)}(x)}{\partial a} = \frac{1}{2\delta} \left(\sigma_{a+\delta}^{(\beta)}(x) - \sigma_{a-\delta}^{(\beta)}(x) \right), \tag{8}$$

$$\frac{\partial S_{a,\delta}^{(\beta)}(x)}{\partial \delta} = \frac{1}{2\delta} \left(\sigma_{a+\delta}^{(\beta)}(x) + \sigma_{a-\delta}^{(\beta)}(x) \right) - \frac{1}{\delta} S_{a,\delta}^{(\beta)}(x). \tag{9}$$

3.3. The error of the approximation

The squashing function approximates the cut function with an error. This error can be defined in many ways. We have chosen the following definition.

Definition 6. Let the approximation error of the squashing function be

$$\varepsilon_\beta = \langle 0 <_\delta (-\delta)\rangle_\beta = \frac{1}{2\delta} \ln \left(\frac{\sigma_\delta^{(-\beta)}(-\delta)}{\sigma_{-\delta}^{(-\beta)}(-\delta)} \right)^{1/\beta},$$

where $\beta > 0$.

Because of the symmetry of the squashing function $\varepsilon_\beta = 1 - \langle 0 <_\delta \delta\rangle_\beta$, see Fig. 9.

The purpose of measuring the approximation error is the following inverse problem: we want to get the corresponding β parameter for a desired ε_β error. We state the following lemma on the relationship between ε_β and β.

Lemma 7. *Let us fix the value of* δ. *The following holds for* ε_β:

$$\varepsilon_\beta < c \cdot \frac{1}{\beta},$$

where $c = \ln 2/2\delta$ *is a constant.*

Proof.

$$\varepsilon_\beta = \frac{1}{2\delta\beta} \ln \left(\frac{1 + e^{\beta(-\delta+\delta)}}{1 + e^{\beta(-\delta-\delta)}} \right) = \frac{1}{2\delta\beta} \ln \left(\frac{2}{1 + e^{-2\delta\beta}} \right)$$

$$= \frac{\ln 2}{2\delta\beta} - \frac{\ln(1 + e^{-2\delta\beta})}{2\delta\beta} < c \cdot \frac{1}{\beta}. \qquad \square$$

So the error of the approximation can be upper bounded by $c \cdot \frac{1}{\beta}$, which means that by increasing parameter β, the error decreases by the same order of magnitude.

4. Approximation of piecewise linear membership functions

In fuzzy theory triangular and trapezoidal membership functions play an important role. For example, fuzzy control uses mainly this type of membership functions because of their easy handling. They are piecewise linear, hence they cannot be continuously differentiated. Our motivation was to construct an approximation which has the same properties in the limit as the approximated membership function and has a continuous gradient. If we are using approximated piecewise linear membership functions in fuzzy control systems then they can be tuned by a gradient-based optimization method and we can get the optimal parameters of the membership functions.

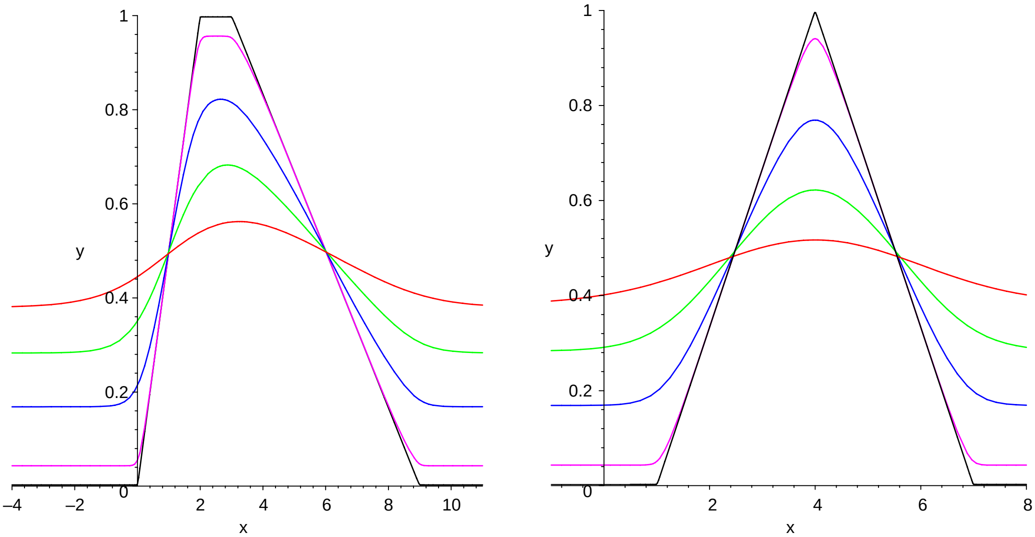

Fig. 10. The approximation of a trapezoid and a triangular membership function.

Piecewise linear membership functions can be constructed by generalized cut functions, so they can be approximated by squashing functions and an appropriate conjunction operator. We have chosen the Łukasiewicz conjunction. But in the conjunction's formula instead of using the cut function we also used the squashing function, so *the membership function and the operator are constructed using the same component*.

To describe a trapezoid membership function using the conjunction operator and two squashing functions four parameters are needed, namely a_1, δ_1 and a_2, δ_2, where a_1 and a_2 give the positions of its left and right sides, and δ_1 and δ_2 give its left and right slopes. The two β parameters of the squashing functions have to have opposite signs to form a trapezoid or triangle, and of course the equations $a_1 < a_2$ and $a_1 + \delta_1 \leqslant a_2 - \delta_2$ must hold.

So the approximation of a trapezoid membership function is the following (see Fig. 10):

$$S^{(\beta)}_{1/2,1/2}\left(S^{(\beta)}_{a_1,\delta_1}(x) + S^{(-\beta)}_{a_2,\delta_2}(x) - 1\right) \tag{10}$$

with pliant notation:

$$\langle\langle a_1 <_{\delta_1} x\rangle_\beta + \langle a_2 <_{\delta_2} x\rangle_{(-\beta)} - 1\rangle_\beta. \tag{11}$$

As a special case of the trapezoid membership function we get the triangular membership function. To describe one, only two parameters are needed, the center a, and its fuzziness δ (see Fig. 10).

Definition 8. Let the approximation of the triangular membership function be defined as (in pliant notation)

$$\begin{aligned}\langle x \sim_\delta a\rangle_\beta = \langle&\langle(a - \delta/2) <_{\delta/2} x\rangle_\beta \\ &+\langle(a + \delta/2) <_{\delta/2} x\rangle_{-\beta} - 1\rangle_\beta,\end{aligned}$$

where a is its center and δ is its fuzziness.

By this way we can represent a fuzzy number by squashing functions.

5. Conclusion

In this paper we approximated piecewise linear membership functions, so that they can be tuned by gradient-based optimizations in a fuzzy control system to achieve better results. We have reviewed the cut function which is the basis of the Łukasiewicz operator class. This cut function is piecewise linear, it cannot be continuously differentiated. We have created an approximation of the cut function (the squashing function) by sigmoid functions with good analytical properties, for example simple derivatives, fast convergence and easy calculation, and applied this approximation to piecewise linear membership functions.

References

[1] R. Ackermann, An Introduction to Many-Valued Logics, Dover, New York, 1967.
[2] A. Bagis, Determining fuzzy membership functions with tabu search—an application to control, Fuzzy Sets and Systems 139 (2003) 209–225.
[3] S. Bodjanova, A generalized α-cut, Fuzzy Sets and Systems 126 (2002) 157–176.
[4] C. Chen, W. Chen, Fuzzy controller design by using neural network techniques, IEEE Fuzzy Systems, Vol. 2.
[5] C.-L. Chen, S.-N. Wang, C.-T. Hsieh, F.-Y. Chang, Theoretical analysis of a fuzzy-logic controller with unequally spaced triangular membership functions, Fuzzy Sets and Systems 101 (1999) 87–108.
[6] P. Chandra, Y. Singh, An activation function adapting training algorithm for sigmoidal feedforward networks, Neurocomputing 61 (1994) 429–437.
[7] R. Cignoli, I.M.L. D'Ottaviano, D. Mundici, Algebraic foundations of many-valued reasoning, Trends in Logic, Vol. 7.
[8] M. Civanlar, H. Trussel, Constructing membership functions using statistical data, Fuzzy Sets and Systems 17 (1986) 1–14.
[9] M. Denna, G. Mauri, A.M. Zanaboni, Learning fuzzy rules with tabu search—an application to control, IEEE Trans. Fuzzy Systems 7 (1999) 295–318.
[10] J. Dombi, Membership function as an evaluation, Fuzzy Sets and Systems 35 (1990) 1–21.
[11] D. Dubois, H. Prade (Eds.), Fundamentals of Fuzzy Sets, Kluwer, Dordrecht, 2000.
[12] L.L.A. Grauel, Construction of differentiable membership functions, Fuzzy Sets and Systems 101 (1999) 219–225.
[13] P. Hájek, Metamathematics of Fuzzy Logic, Kluwer, Dordrecht, 1998.
[14] D. Karaboga, Design of fuzzy logic controllers using tabu search algorithms, Biennial Conf. North American Fuzzy Inf. Processing Society, 1996, pp. 489–491.
[15] M. Kodaka, Requirements for generating sigmoidal timecourse aggregation in nucleation-dependent polymerization model, Biophys. Chem. 107 (2004) 243–253.
[16] J.T.B. Thomas, A. Runkler, Function approximation with polynomial membership functions and alternating cluster information, Fuzzy Sets and Systems 101 (1999) 207–218.