ELSEVIER

Fuzzy Sets and Systems 125 (2002) 105–119

FUZZY sets and systems

www.elsevier.com/locate/fss

Stability of interpolative fuzzy KH controllers ☆

Domonkos Tikk[a,*,1], István Joó[a], László T. Kóczy[a], Péter Várlaki[b], Bernhard Moser[c], Tamás D. Gedeon[d,1]

[a] Department of Telecommunications & Telematics, Budapest University of Technology and Economics, H-1117 Pázmány sétány 1/d, Budapest, Hungary

[b] Department of Road Vehicles, Budapest University of Technology and Economics, H-1111 Sztoczek u. 6., Budapest, Hungary

[c] Fuzzy Logic Laboratory Linz-Hagenberg, Johannes Kepler University, A-4040, Linz, Austria

[d] School of Information Technology, Murdoch University, South Street, Murdoch, 6150 WA, Australia

Received 15 February 1999; received in revised form 18 April 2000; accepted 23 May 2000

Abstract

The classical approaches in fuzzy control (Zadeh and Mamdani) deal with dense rule bases. When this is not the case, i.e. in sparse rule bases, one has to choose another method. Fuzzy rule interpolation [proposed first by Kóczy and Hirota, Internat. J. Approx. Reason. 9 (1993) 197–225] offers a possibility to construct fuzzy controllers (KH controllers) under such conditions. The main result of this paper shows that the KH interpolation method is stable. It also contributes to the application oriented use of Balázs–Shepard interpolation operators investigated extensively by researchers in approximation theory. The numerical analysis aspect of the result contributes to the well-known problem of finding a stable interpolation method in the following sense. We would like to approximate a function that is only known at distinct points (e.g. where it can be measured). Since measurement errors cannot be eliminated in practice, it is a natural requirement that the interpolation method should be stable independently from the selection of measurement points. The classical interpolation methods generally do not fulfil this condition, only with certain strong restrictions concerning the measurement points. If we neglect the classical form of the approximating function (polynomial and trigonometrical), we can produce better behaving approximation. The KH controller approximating function, being a simple fractional function, fulfils the stability condition. This can also be interpreted as KH controllers are universal approximators in the space of continuous functions (with respect to, e.g., the supremum norm). © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Fuzzy inference systems; Fuzzy control; Fuzzy rule interpolation; Stability; Universal approximation

☆ Research supported by the Hungarian Scientific Research Fund (OTKA) Grant No. D034614.

* Corresponding author.

E-mail address: tikk@ttt.bme.hu (D. Tikk).

[1] Partially funded by the Australian Research Council large grants scheme.

0165-0114/02/$ - see front matter
PII: S0165-0114(00)00104-4

1. Introduction

The classical approaches of fuzzy control deal with dense rule bases where the input universe of discourse is fully α-covered by the antecedent fuzzy sets of the rule base in each dimension. (Fully α-covering means that for any $x=(x_1,x_2,\ldots,x_N)\in X=X_1\times X_2\times\cdots\times X_N$ in each X_i $(i=1,\ldots,N)$ there is a fuzzy set $A\in\mathscr{F}(X_i)$ for that $A(x_i)\geqslant\alpha$ holds, where $\alpha\in(0,1]$, usually $\alpha\geqslant 0.5$. Here, $\mathscr{F}(Z)$ denotes the set of fuzzy sets defined on Z.) However, this is not always the case, because several reasons can lead to α-sparse or sparse rule bases (i.e. where at least one of the input universes is not fully (α-)covered by the rule antecedents). A good example is when the starting term set is of α-cover type. Nevertheless, by tuning the rules, the antecedents are partially shifted and shrunk so the tuned model will contain gaps [6]. Another cause can be the omission of rules from a complete but redundant rule base in order to reduce complexity, which may also result easily in a sparse rule base [17]. In such a sparse rule base no rule can be applied, and hence, no consequent can be constructed by means of the classical methods if the observation is within a gap. The fuzzy rule interpolation method (proposed first by Kóczy and Hirota [15]) provides a tool to construct fuzzy controllers handling sparse rule bases.

On the other hand, there is a great demand for finding a stable interpolating method among researchers in the field of mathematical analysis. In general, the interpolation problem aims at the exact implementation (or approximation) of a continuous function that is known only at certain points (i.e. where it can be measured). It is a natural requirement of an interpolation method that it should be stable independently from the measurement points, in other words, it should converge to the approximated function regardless of the measurement points if their number converges to infinity. The classical mathematical interpolation methods generally do not fulfil this condition. These approaches require that the approximated function should be measured at some special exact points. In practice, it can hardly be solved because of the uncertainty of the measuring process. If we consider the input–output function of the KH interpolation method as a function approximator, it fulfils the stability condition, that is it converges to the approximated function independently from the points where the latter have been measured. Further, a class of stable interpolating functions can be derived from the original KH function. We also present the so-called (Balázs–)Shepard interpolation which has a strong relation with the KH interpolation for every fixed α level. Some corresponding results for Shepard interpolation can be exploited to characterize the approximation behaviour of the KH interpolation.

This paper is organized as follows. Section 2 gives a short overview of the KH rule interpolation method. Section 3 recalls some problems from the field of mathematical (numerical) analysis relating to the selection of the measurement points. Section 4 reviews the relating results about Shepard interpolation in approximation theory. Section 5 introduces the main results concerning the stability property of the KH interpolation. At the end of this section, a counterexample is introduced to illustrate that the second statement of the main theorem is sensitive to the value of its parameters. Finally, Section 6 presents our conclusions.

2. Rule interpolation

Rule bases containing gaps require completely new techniques of reasoning and control. This family of methods works well only if the system has some “nice” properties: it is not allowed to behave too unexpectedly in the areas where the model does not cover. Luckily, in practice such a nice behaviour might be expected in most cases. The term for the class of systems where the following algorithms are applicable is *interpolative system*. (We remark here that interpolativity is connected with the Shannon sampling theorem.)

Let us introduce the extended concept of rule interpolation. The starting ideas are the Extension and Resolution Principles. The former states that the solution of a problem for fuzzy sets can be obtained by solving it first for arbitrary α-cuts (N.B. these are crisp sets) and then extending the solution to the fuzzy case. The latter describes the decomposition of fuzzy sets to α-cuts:

$$A = \bigcup_{\alpha \in [0,1]} \alpha A_\alpha \tag{1}$$

(here the union means maximum).

Every fuzzy set can be approximated by the family of approximations of its cuts. Although theoretically an infinite number of cuts should be treated separately, in most practical cases, if the membership functions are piecewise linear, it is often sufficient to calculate for only a few important or typical cuts [18,25–27], e.g. $\alpha = 0$ and 1.

Some conditions must be fulfilled for the applicability of the interpolation method: the fuzzy sets must be normal and convex (briefly: CNF sets), and the state variables, including X_i and Y as well, must be bounded and gradual which guarantees that a full ordering on each of them exists. This implies the existence of a partial ordering on the whole input universe, therefore we can introduce the following partial ordering among the elements of X (i.e. among CNF sets) with the help of their α-cuts. If

$$\forall \alpha \in [0,1]\colon\ \inf\{A_\alpha\} \leqslant \inf\{B_\alpha\} \quad \text{and} \quad \sup\{A_\alpha\} \leqslant \sup\{B_\alpha\},$$

then A and B are comparable, i.e. $A \prec B$.

Among comparable fuzzy sets the new concept of distance can be defined. For all $\alpha \in [0,1]$, in practice, only for each significant α (e.g. $\alpha = 0$ and 1), the two extremes of the α-cuts are considered and their pairwise distances will be named the *lower* and the *upper fuzzy distance* of the two cuts. The definition of these distances results in two families of distances as follows:

$$\mu_{d_{\mathrm{L}}(A,B)}(z_i) = \sum_{\alpha \in [0,1]} \alpha / |\inf\{A_\alpha\} - \inf\{B_\alpha\}|,$$

$$\mu_{d_{\mathrm{U}}(A,B)}(z_i) = \sum_{\alpha \in [0,1]} \alpha / |\sup\{A_\alpha\} - \sup\{B_\alpha\}|,$$

where $z_i \in Z$ is the variable representing the possible values of distances in X_i and $\sum \alpha/x_i$ denotes $\mu(x_i) = \alpha$. (For more details on these distances see [16].) Using the concept of fuzzy distance, the closeness of two comparable fuzzy sets can be determined even if their supports are disjoint.

With the help of fuzzy distance, the classical methods of function approximation can be applied to rule bases in order to obtain a proper consequent, even if they are sparse. Using (1), a rule base can be represented by a family of hyperintervals in $X \times Y$, that is, for every α, $A_{i\alpha}$, $B_{i\alpha}$ form a hyperinterval for CNF sets. By the Extension Principle, instead of approximating the original fuzzy mapping, only its *important cuts* (level set) will be approximated. Hence, the problem is reduced to a family of non-fuzzy approximation problems which can be solved with any of the classical function approximation methods like interpolation or extrapolation, etc.

The simplest of these methods is the linear interpolation of two rules for the area between their antecedents. This can be applied if the observation is located so that

$$A_{i1} \prec A^* \prec A_{i2} \quad \text{and} \quad B_{i1} \prec B_{i2}.$$

Using the concept of fuzzy distance, the following fundamental equation of *linear interpolation* can be written, in accordance with the gradual semantic interpretation of fuzzy rules by [13]

$$d(A^*, A_{i1}) : d(A^*, A_{i2}) = d(B^*, B_{i1}) : d(B^*, B_{i2}). \tag{2}$$

After decomposing (2) to every $\alpha\in[0,1]$ of the level set, it can be solved for B^*_α. The following formulae are the solution for linear KH controllers:

$$\min\{B^*_\alpha\}=\frac{\dfrac{\inf\{B_{i1\alpha}\}}{d_{\mathrm{L}}(A^*_\alpha,A_{i1\alpha})}+\dfrac{\inf\{B_{i2\alpha}\}}{d_{\mathrm{L}}(A^*_\alpha,A_{i2\alpha})}}{\dfrac{1}{d_{\mathrm{L}}(A^*_\alpha,A_{i1\alpha})}+\dfrac{1}{d_{\mathrm{L}}(A^*_\alpha,A_{i2\alpha})}};$$

$$\max\{B^*_\alpha\}=\frac{\dfrac{\sup\{B_{i1\alpha}\}}{d_{\mathrm{U}}(A^*_\alpha,A_{i1\alpha})}+\dfrac{\sup\{B_{i2\alpha}\}}{d_{\mathrm{U}}(A^*_\alpha,A_{i2\alpha})}}{\dfrac{1}{d_{\mathrm{U}}(A^*_\alpha,A_{i1\alpha})}+\dfrac{1}{d_{\mathrm{U}}(A^*_\alpha,A_{i2\alpha})}}. \tag{3}$$

For the two families of solutions to determine a fuzzy set B^*, they should satisfy $\min\{B^*_\alpha\}\leqslant\max\{B^*_\alpha\}$ for every α, cf. [16]. In certain cases this condition is not satisfied, and hence we obtain a conclusion not directly interpretable as a fuzzy set. To solve this problem alternative methods were proposed which provided a fuzzy set as a conclusion regardless of the shape of the membership functions involved [2,31].

The principle of interpolating two rules can be extended in many different ways. The most obvious extension of the linear interpolation is the interpolation of more rules flanking the observation (with respect to the ordering on the input space) from both sides. Intuitively, the further a rule from the observation is located, the less weight the respective consequent plays in the construction of the conclusion. The general KH interpolation (also termed extended) for this type of interpolation are obtained from the solution of (3) repeatedly for the points taken into consideration and then normalizing the solution as

$$\min B^*_\alpha=\frac{\sum_{i=1}^{m}\dfrac{\inf\{B_{i\alpha}\}}{d_{\mathrm{L}}(A^*_\alpha,A_{i\alpha})}}{\sum_{i=1}^{m}\dfrac{1}{d_{\mathrm{L}}(A^*_\alpha,A_{i\alpha})}},\qquad \max B^*_\alpha=\frac{\sum_{i=1}^{m}\dfrac{\sup\{B_{i\alpha}\}}{d_{\mathrm{U}}(A^*_\alpha,A_{i\alpha})}}{\sum_{i=1}^{m}\dfrac{1}{d_{\mathrm{U}}(A^*_\alpha,A_{i\alpha})}}, \tag{4}$$

where m is the number of rules taken into account. More details on this method can be found in [15].

In the main part of this work (Section 5) we will use a slightly different version of the general KH interpolation. So, we change (4) by taking the Nth power of the distance function d, where N denotes the dimension of the domain X:

$$\min B^*_\alpha=\frac{\sum_{i=1}^{m}\dfrac{\inf\{B_{i\alpha}\}}{d^N_{\mathrm{L}}(A^*_\alpha,A_{i\alpha})}}{\sum_{i=1}^{m}\dfrac{1}{d^N_{\mathrm{L}}(A^*_\alpha,A_{i\alpha})}},\qquad \max B^*_\alpha=\frac{\sum_{i=1}^{m}\dfrac{\sup\{B_{i\alpha}\}}{d^N_{\mathrm{U}}(A^*_\alpha,A_{i\alpha})}}{\sum_{i=1}^{m}\dfrac{1}{d^N_{\mathrm{U}}(A^*_\alpha,A_{i\alpha})}}, \tag{5}$$

We will refer to formula (5) as stabilized (general) KH interpolation. Its stability will be proved in Section 5.

3. Problems of measurement points selection

Interpolation is a method where a continuous function is approximated by some analytical (e.g. algebraic) function, based on a finite number of known points of the original function. The essential idea is to construct a polynomial which coincides at certain given points (the so-called *measurement points*) with the function to be approximated, $f(x)$.

The selection of the measurement points plays an important role in the theory of interpolation. This section briefly reviews the related results (cf. [21]).

The first result in this direction was given by Gauss. The quadrature formula

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{k=1}^{n} A_k f(x_k) \tag{6}$$

is exact if the degree of polynomial $f(x)$ is at most $n-1$ with arbitrary selection of the measurement points x_i. This holds because in this case, $f(x)$ is equal to its own Lagrange interpolation polynomial. (Here we approximate the value of the "determined integral".) Gauss showed that if the measurement points were selected specially, then (6) was exact for any polynomial of degree at most $2n-1$. This also holds for the following more general case.

Let $p(x)$ be a weighting function on $[a,b]$. Let us construct the Lagrange polynomial $L(x) = \sum_{k=1}^{n} f(x_k) l_k(x)$ for the continuous $f(x)$ on $[a,b]$, where $x_1,\dots,x_n$ are the measurement points and $l_k(x)$, $1 \leqslant k \leqslant n$ are the Lagrange coefficients. Multiplying $L(x)$ by $p(x)$ and then integrating we get

$$\int_a^b p(x)L(x)\,\mathrm{d}x = \sum_{k=1}^{n}\left(\int_a^b p(x)l_k(x)\,\mathrm{d}x\right) f(x_k) = \sum_{k=1}^{n} A_k f(x_k),$$

and if $L(x)$ is considered as the approximation of $f(x)$ then we obtain the

$$\int_a^b p(x)f(x)\,\mathrm{d}x = \sum_{k=1}^{n} A_k f(x_k) \tag{7}$$

approximating quadrature formula, this is exact for any polynomial $f(x)$ of degree at most $2n-1$ with the special selection of measurement points described in the next paragraph. (When $p(x)=1$, the formulae (6) and (7) are equivalent.)

According to Gauss, to achieve the exact interpolation, the measurement points have to satisfy the following condition. If the measurement points are chosen as the zeros of a polynomial $\omega_n(x)$ of degree n which is orthogonal with respect to $p(x)$ for every polynomial of degree at most $n-1$, then the formula (7) is exact.

Example 1. If $a=-1$, $b=1$ and $p(x)=1$, then the solution of the Gauss' problem is formed by the zeros of the Legendre polynomial $X_n(x)$.

Let us consider a more general interpolation problem, the so-called Hermite interpolation. The aim is to construct the polynomial $H(x)$ of the minimal degree which satisfies the conditions

$$H^{(r)}(x_i) = y_i^{(r)} \quad (i = 1,2,\dots,n;\ r = 0,1,\dots,a_i - 1), \tag{8}$$

where $y_i^{(r)}$ and x_i are given, and r is the highest rank of derivatives taken into account. We consider only the case when $a_1 = a_2 = \cdots = a_n = 2$. The solution of (8) is given by the formula

$$H(x) = \sum_{k=1}\left[1 - \frac{\omega''(x_k)}{\omega'(x_k)}(x - x_k)\right] l_k^2(x) + \sum_{k=1}^{n} y_k'(x - x_k) l_k^2(x) \tag{9}$$

using the ordinary notation. We note that the degree of $H(x)$ is $2n-1$.

Usually, the continuity of the function and of its derivatives does not give sufficient conditions for the convergence of $H(x)$ Numerous counterexamples are known, e.g. from Bernstein or Marcinkiewicz, cf. [21]. Nevertheless, Fejér has shown if x_k are the zeros of the Tchebysheff polynomial, the convergence can be proved [21].

As we have shown in this section the choice of the measurement points plays a very important role in many areas concerning numerical interpolation. From the practical point of view, if we would approximate a function by interpolation, measuring its value at distinct points, it is almost impossible to choose them exactly as, e.g., the zeros of the Tchebysheff polynomial.

The interpolation method introduced by Kóczy and Hirota [15] for fuzzy control satisfies the property of convergence independently from the values of x_k as will be shown in Section 5.

4. Shepard interpolation and KH interpolation

The Shepard interpolation method was first introduced in [24] for arbitrarily placed bivariate data as

$$S_0(f,x,y)=\begin{cases} f_k & \text{if } (x,y)=(x_k,y_k) \text{ for some } k,\\ \left(\sum\limits_{k=0}^{n} f(x_k,y_k)/d_k^{\lambda}\right)\Big/\left(\sum\limits_{k=0}^{n} 1/d_k^{\lambda}\right) & \text{otherwise,}\end{cases} \tag{10}$$

where measurement points x_k, y_k ($k\in[0,n]$) are irregularly spaced on the domain of $f\in\mathbb{R}^2\to\mathbb{R}$, $\lambda>0$, and $d_k=[(x-x_k)^2+(y-y_k)^2]^{1/2}$. This function can be used typically when a surface model is required to interpolate scattered spatial measurements. This problem is encountered in such areas as pattern recognition, geology, cartography, earth sciences, fluid dynamics and many others.

Because this method has some unsatisfactory properties (mostly for small $\lambda\in(0,1]$), in [4] an extension of Shepard's method was introduced which eliminated these shortcomings.

From a practical point of view, formula (10) also has the defect that if another interpolation point is added, then all weights must be reformulated. This deficiency can be eliminated by the recursive "expandable" Shepard method proposed also in [4]. This modification is based on the similar idea of how one can derive the Newton form of linear interpolation from the Lagrange-type interpolation.

Beside the application oriented investigation of Shepard's method such as [4] (see also [3,14,20,23,22,5]), an increasing interest has arisen from mathematical researchers to examine the approximation property of formula (10).

Formula (10) was generalized as follows:

$$S_{n,\lambda}(f,x)=\frac{\sum_{k=0}^{n} f(x_k)(x-x_k)^{-\lambda}}{\sum_{k=0}^{n}(x-x_k)^{-\lambda}},\quad \lambda>0,\ n=1,2,\ldots \tag{11}$$

for an arbitrary $f\in C[0,1]$, where x_k ($k=0,\ldots,n$), in general, denotes the nodes of the equidistant distribution of the domain $[0,1]$. We note that fixing the domain to the interval $[0,1]$ does not mean any restriction, because by a proper linear transformation any finite interval can be mapped into another one.

The use of (11) type of rational functions as approximating means was first discovered by Balázs [1]. (After his name this operator is often termed in the literature of approximation theory as Balázs–Shepard operator.) The properties of the operator (11) were widely investigated by Szabados [29,30], Somorjai [28], by several Italian mathematicians: Della Vecchia, Mastroianni, Criscuolo, etc., see e.g. [9–12].

The first result on the approximation property of the Shepard operator was published by Szabados [29], where the author showed that for $\lambda=4$ the operator (11) realized the so-called Jackson order of approximation, i.e.,

$$\|f-S_n(f)\|=\mathrm{O}(\omega(f,n^{-1})),$$

where $\omega(f,n^{-1})$ was the modulus of smoothness (or continuity) of function f, that is

$$\omega(f,n^{-1}) = \max_{\substack{x,y\in\bar{\Omega}\\ |x-y|\leqslant n^{-1}}} |f(x)-f(y)|. \tag{12}$$

The saturation problem, i.e., the determination of the optimal order of convergence was also investigated by Somorjai [28]. He gave a complete description of the order and class of saturation when the weights $(x-x_k)^{-\lambda}$ were replaced by $|x-k/n|^{-\lambda}$, $k=0,\dots,n$ for $\lambda>2$. In [28] it was shown that for $\lambda>2$

$$\|f-S_n(f,x)\| = \begin{cases} \mathrm{o}(1/n) & \text{iff } f=\text{const.},\\ \mathrm{O}(1/n) & \text{iff } f\in \mathrm{Lip}1. \end{cases}$$

The function $f:[a,b]\to\mathbb{R}$ is called Lipschitz continuous with Lipschitz coefficient α (notation: $f\in\mathrm{Lip}\,\alpha$) if

$$|f(x)-f(y)|\leqslant\alpha|x-y| \quad \text{for all } x,y\in[a,b]. \tag{13}$$

In [12] the authors investigated the saturation problem of (11) for $\lambda=2$. They showed that if

$$\|S_{n,2}(f,x)-f(x)\| = \mathrm{o}(n^{-1})$$

is satisfied for a continuous function f for all $x\in[0,1]$, then f is constant. This statement shows that the $\lambda=2$ case represents additional difficulties compared to $\lambda>2$, because we do not have strong localization like for $\lambda>2$. This peculiarity can be best illustrated by noting that the linear functions are not in the saturation class of the operator $S_{n,2}$. In [12] it is also shown that the saturation class of $S_{n,2}$ is the set of constant functions. Furthermore, non-constant linear functions do not belong to the saturation class, and the saturation class is contained in $\bigcup_{\alpha<1}\mathrm{Lip}\,\alpha$. Here it is reasonable to remark that classical approximation processes (e.g. Bernstein polynomials) are all saturated with order n^{-1}, and with trivial classes of functions, either the set of linear or the set of constant functions.

In [30] direct and converse estimates were given for the case where the weights are replaced by their absolute values $|x-k/n|^{-\lambda}$ (with equidistant nodes) for all $\lambda\geqslant1$. Namely, it was shown that

$$\|S_{n,\lambda}(f,x)-f(x)\| = \begin{cases} \mathrm{O}\left(\omega\left(f,\dfrac{1}{n}\right)\right) & \text{if } \lambda>2,\\ \mathrm{O}(n^{1-\lambda})\displaystyle\int_{1/n}^{1} t^{-\lambda}\omega(f,t)\,\mathrm{d}t & \text{if } 1<\lambda\leqslant2,\\ \mathrm{O}(\log^{-1} n)\displaystyle\int_{1/n}^{1} t^{-1}\omega(f,t)\,\mathrm{d}t & \text{if } \lambda=1 \end{cases} \tag{14}$$

for any $f\in C[0,1]$. These results do not give the optimal order of approximation for $\lambda<2$, thus in this paper the saturation problem remains unsettled.

In [10] the authors obtained pointwise simultaneous approximation estimates for rational operators, being a generalization of the Shepard operator (11), which were not possible by algebraic polynomials. In this result, opposed to the previous ones, the nodes are placed arbitrarily.

In [8] estimation of the convergence of (11) was given with respect to different matrices of knots. For some special knot point matrices, such as zeros of orthogonal polynomials, the F-stable property (stability in Fejér sense, see also, e.g. [23]) was also proved.

In [11] the characterization of Shepard operator on infinite intervals was given.

Summarizing the characterization of the interpolatory Shepard operator, we can state that one of the main goals, the saturation problem, has been solved only for the case $\lambda>2$ and for equispaced nodes. Other node point systems were also investigated, but many open problems remained unsettled.

It was shown in this section the Shepard operator was extensively investigated by researchers in the field of approximation theory and by authors interested in some applications, such as fitting data, curves, and surface, fluid dynamics problems. In the next section, we will return to Shepard interpolation showing its relationship with KH interpolation.

5. Main results

Definition 2. Let $\mathbb{R}^N \supset \Omega = [a_1, b_1] \times \cdots \times [a_N, b_N]$, further let $\{\Gamma_n\}_{n=1}^{\infty}$ be a sequence of finite subsets of Ω with $\#\Gamma_n = n$. If

$$\forall \varepsilon > 0\ \exists n_0\ \forall \omega \in \Omega\ \forall n \geqslant n_0\colon\ \left| \frac{\#(\Gamma_n \cap \omega)}{\#\Gamma_n} - \frac{|\omega|}{|\Omega|} \right| < \varepsilon, \tag{15}$$

then the set Γ_n is uniformly distributed on the domain Ω. Here $\#(\Gamma_n \cap \omega)$ denotes the cardinality of the finite set $(\Gamma_n \cap \omega)$ and $|\omega|$ is the Lebesque measure of ω.

Theorem 3. *Consider the L_p norm $\|\cdot\|_p$ with $p \in [1, \infty]$, the domain $\mathbb{R}^N \supset \Omega = [a_1, b_1] \times \cdots \times [a_N, b_N]$, and a continuous function $f : \Omega \to \mathbb{R}$, then*

$$\forall x \in \Omega\colon\ \lim_{n\to\infty} K_n(f, x) := \lim_{n\to\infty} \sum_{k=1}^{n} f(x_k^{(n)}) \frac{\dfrac{1}{\|x - x_k^{(n)}\|_p^N}}{\displaystyle\sum_{j=1}^{n} \frac{1}{\|x - x_j^{(n)}\|_p^N}} = f(x), \tag{16}$$

where measurement points $x_k^{(n)}$ are uniformly distributed on Ω in the sense of (15).

Proof. Let us estimate the following difference:

$$K_n(f, x) - f(x) = \sum_{k=1}^{n} [f(x_k) - f(x)] \frac{\dfrac{1}{\|x - x_k\|_p^N}}{\displaystyle\sum_{j=1}^{n} \frac{1}{\|x - x_j\|_p^N}}.$$

For brevity, we have written x_k instead of $x_k^{(n)}$. For arbitrary $\varepsilon > 0$, because of the continuity of f, we can choose a sufficiently small real number $\delta > 0$, such that for $\|x - x_k\|_p < \delta$

$$|f(x_k) - f(x)| < \varepsilon \tag{17}$$

holds. δ can be chosen on $\bar{\Omega}$ since f is uniformly continuous. Then for fixed x, the numbers $1, 2, \ldots, n$ are divided into two groups: group I contains those numbers k for which $\|x - x_k\|_p \leqslant \delta$, and group II contains the remaining ones

$$\sum_{k\in \mathrm{I}} [f(x_k) - f(x)] \frac{\dfrac{1}{\|x - x_k\|_p^N}}{\displaystyle\sum_{j=1}^{n} \frac{1}{\|x - x_j\|_p^N}} + \sum_{k\in \mathrm{II}} [f(x_k) - f(x)] \frac{\dfrac{1}{\|x - x_k\|_p^N}}{\displaystyle\sum_{j=1}^{n} \frac{1}{\|x - x_j\|_p^N}} =: \Sigma_{\mathrm{I}} + \Sigma_{\mathrm{II}}. \tag{18}$$

Because of (17)

$$|\Sigma_{\rm I}| \leqslant \varepsilon \sum_{k\in{\rm I}} \frac{\dfrac{1}{\|x-x_k\|_p^N}}{\sum\limits_{j=1}^{n} \dfrac{1}{\|x-x_j\|_p^N}} \leqslant \varepsilon \sum_{k=1}^{n} \frac{\dfrac{1}{\|x-x_k\|_p^N}}{\sum\limits_{j=1}^{n} \dfrac{1}{\|x-x_j\|_p^N}} = \varepsilon. \tag{19}$$

Now let us turn to estimate $\Sigma_{\rm II}$. The continuous function f is bounded on the compact domain $\bar{\Omega}$, thus

$$|f(x_k) - f(x)| \leqslant 2 \max_{\bar{\Omega}} |f| = 2\|f\|_C =: M$$

and so

$$|\Sigma_{\rm II}| \leqslant \sum_{k\in{\rm II}} |f(x_k) - f(x)| \frac{\dfrac{1}{\|x-x_k\|_p^N}}{\sum\limits_{j=1}^{n} \dfrac{1}{\|x-x_j\|_p^N}} \leqslant M \sum_{k\in{\rm II}} \frac{\dfrac{1}{\|x-x_k\|_p^N}}{\sum\limits_{j=1}^{n} \dfrac{1}{\|x-x_j\|_p^N}} \leqslant M \frac{V}{V+W}, \tag{20}$$

where $V = \sum_{j\in{\rm II}} 1/(\|x-x_j\|_p^N)$, and $W = \sum_{j\in{\rm I}} 1/(\|x-x_j\|_p^N)$. In the following, without loss of generality, we will apply the maximum norm $\|\cdot\| = \|\cdot\|_C$ to simplify further computation. To estimate the sums V and W we proceed as follows. We define a hypercube covering the domain of f. Then we estimate the sums by counting the number of grid points on the surface of hypercubes (more precisely on the shell of hypercubes) in terms of the length of edges. Finally, summing these values on all the shells that forms the covering hypercube we obtain the required estimations.

Assuming uniform distribution (15) of knot points for sufficiently great n, this results in an equidistant N-dimensional grid on Ω. Choose a sufficiently great hypercube $\mathscr{H}(R)$ with edge R to cover the entire domain of function f. Let η be the grid spacing in the hypercube $\mathscr{H}(R)$ (distance between neighbouring nodes)

$$\eta = \frac{R}{\sqrt[N]{n}}. \tag{21}$$

Further, let us compute the asymptotic number of the grid points on the surface of the hypercube $\mathscr{H}(r)$

$$\frac{2N(r)^{N-1}}{\eta^{N-1}} = 2N\left(\frac{r}{\eta}\right)^{N-1} = 2N\left(\frac{r\sqrt[N]{n}}{R}\right)^{N-1} = 2Nn^{1-1/N}\left(\frac{r}{R}\right)^{N-1}.$$

First, let us estimate V by summing the reciprocal of the distances on the surface of hypercubes $\mathscr{H}(r_i)$ where $r_{i+1} - r_i = \eta$, and $\delta \leqslant r_i \leqslant R$:

$$V = \sum_{i=0}^{i_{\max}^V} 2Nn^{1-1/N}\left(\frac{r_i}{R}\right)^{N-1} \frac{1}{r_i^N}, \tag{22}$$

where

$$i_{\max}^V = \frac{R-\delta}{\eta} = \frac{R-\delta}{R}\sqrt[N]{n}$$

from $r_i = \delta + i\eta = \delta + iR/\sqrt[N]{n} \leqslant R$. Hence,

$$V = 2Nn^{1-1/N}\frac{1}{R^{N-1}}\sum_{i=0}^{i_{\max}^V}\frac{1}{r_i} = 2Nn^{1-1/N}\frac{1}{R^{N-1}}\sum_{i=0}^{i_{\max}^V}\frac{1}{\left(\frac{R}{\sqrt[N]{n}}\right)\left(\delta\frac{\sqrt[N]{n}}{R}+i\right)}$$

$$= 2Nn\frac{1}{R^N}\sum_{i=0}^{(R-\delta)/\eta}\frac{1}{\left(\delta\frac{1}{\eta}+i\right)}. \tag{23}$$

We can estimate W analogously, only the boundaries of the sum have to be modified. More precisely, there are two cases. If x identical to a grid point, W takes the value ∞, thus (16) is proved by (18)–(20). Otherwise there is a closest grid point x_j to x such that

$$\delta_0 := \|x - x_j\| = \min_{1\leqslant k\leqslant n}\|x - x_k\| < \eta.$$

Thus,

$$W = \sum_{k,\delta_0\leqslant\|x-x_k\|\leqslant\delta}\frac{1}{\|x-x_k\|^N} = 2Nn\frac{1}{R^N}\sum_{i=0}^{i_{\max}^W}\frac{1}{\left(\delta_0\frac{1}{\eta}+i\right)} \tag{24}$$

and here

$$i_{\max}^W = \frac{\delta-\delta_0}{\eta} = \left(\frac{\delta-\delta_0}{R}\right)\sqrt[N]{n}$$

by $r_i = \delta_0 + i\eta = \delta_0 + iR/\sqrt[N]{n} \leqslant \delta$.

By (23) and (24) inequality (20) leads to

$$|\Sigma_{\mathrm{II}}| \leqslant M\frac{2N\frac{n}{R^N}\sum_{i=0}^{i_{\max}^V}\frac{1}{\alpha_n^V+i}}{2N\frac{n}{R^N}\sum_{i=0}^{i_{\max}^V}\frac{1}{\alpha_n^V+i}+2N\frac{n}{R^N}\sum_{i=0}^{i_{\max}^W}\frac{1}{\alpha_n^W+i}} = M\frac{\sum_{i=0}^{i_{\max}^V}\frac{1}{\alpha_n^V+i}}{\sum_{i=0}^{i_{\max}^V}\frac{1}{\alpha_n^V+i}+\sum_{i=0}^{i_{\max}^W}\frac{1}{\alpha_n^W+i}}, \tag{25}$$

where α_n^V and α_n^W are $\delta\sqrt[N]{n}/R$ and $\delta_0\sqrt[N]{n}/R$, respectively.

We use integral approximation for estimating the sums in (25):

$$\sum_{i=0}^{i_{\max}^V}\frac{1}{\alpha_n^V+i} \geqslant \int_{\alpha_n^V+1}^{\alpha_n^V+i_{\max}^V+1}\frac{1}{x+1}\,\mathrm{d}x = \log(\alpha_n^V+i_{\max}^V+2) - \log(\alpha_n^V+2)$$

$$= \log\left(\frac{\frac{\delta}{\eta}+\frac{R-\delta}{\eta}+2}{\frac{\delta}{\eta}+2}\right) \overset{n\to\infty}{=} \log\left(\frac{R}{\delta}\right), \tag{26}$$

analogously,

$$\sum_{i=0}^{i^W_{\max}} \frac{1}{\alpha_n^W + i} \geqslant \log\left(\frac{\dfrac{\delta_0}{\eta} + \dfrac{\delta-\delta_0}{\eta} + 2}{\dfrac{\delta_0}{\eta} + 2}\right) \stackrel{n\to\infty}{=} \log\left(\frac{\delta}{\delta_0}\right). \tag{27}$$

By (25)–(27)

$$|\Sigma_{\mathrm{II}}| \leqslant M \frac{\ln(R/\delta)}{\ln(R/\delta) + \ln(\delta/\delta_0)} = \text{const}\,\frac{\ln R - \ln \delta}{\ln R - \ln \delta_0} < \varepsilon, \tag{28}$$

because for sufficiently great n when $0<\delta_0 \ll \delta < 1 \ll R$ holds, the right-hand side of inequality (28) can be less than arbitrary $\varepsilon > 0$. Observe that δ_0 is independent from δ due to its definition. This completes the proof. □

We can generalize the above theorem as follows:

Theorem 4. *Consider the same conditions as in Theorem* 3. *If we substitute λ for the exponent N in the expression of the distance, the resulting $K_n^\lambda(f,x)$ converges to the approximated function f, as well*:

$$\forall x \in \Omega\colon \lim_{n\to\infty} K_n^\lambda(f,x) := \lim_{n\to\infty} \sum_{k=1}^{n} f(x_k^{(n)}) \frac{\dfrac{1}{\|x - x_k^{(n)}\|_p^\lambda}}{\displaystyle\sum_{j=1}^{n} \frac{1}{\|x - x_k^{(n)}\|_p^\lambda}} = f(x), \tag{29}$$

where ($\lambda \geqslant N$), *and with uniform distribution of measurement points $x_k^{(n)}$*.

Proof. For proving the convergence of (29) we use similar reasoning as above. Let $t=\lambda - N > 0$. Following the train of thought of the former proof and using its notation we should only modify the estimation of V and W (for clarity, here we denote them by V_λ and W_λ, respectively) as follows. Eq. (22) changes as

$$V_\lambda = \sum_{i=0}^{i^V_{\max}} 2Nn^{1-1/N}\left(\frac{r_i}{R}\right)^{N-1} \frac{1}{r_i^\lambda},$$

and we obtain

$$V_\lambda = 2Nn^{N-1/N-\lambda-1} \frac{1}{R^\lambda} \sum_{i=0}^{i^V_{\max}} \left(\frac{\delta}{\eta} + i\right)^{N-\lambda-1} \tag{30}$$

in the same way as in (23). Similarly,

$$W_\lambda = 2Nn^{N-1/N-\lambda-1} \frac{1}{R^\lambda} \sum_{i=0}^{i^W_{\max}} \left(\frac{\delta_0}{\eta} + i\right)^{N-\lambda-1}, \tag{31}$$

hence by (30) and (31)

$$|\Sigma_{\mathrm{II}}| \leqslant M \frac{V_\lambda}{V_\lambda + W_\lambda} = M \frac{\sum_{i=0}^{i^V_{\max}} (\alpha_n^V + i)^{N-\lambda-1}}{\sum_{i=0}^{i^V_{\max}} (\alpha_n^V + i)^{N-\lambda-1} + \sum_{i=0}^{i^W_{\max}} (\alpha_n^W + i)^{N-\lambda-1}}. \tag{32}$$

By integral approximation of the sums in (32)

$$\sum_{i=0}^{i_{\max}^V} \frac{1}{(\alpha_n^V+i)^{t+1}} \geqslant \int_{\alpha_n^V+1}^{\alpha_n^V+i_{\max}^V+1} \frac{1}{(x+1)^{t+1}}\,\mathrm{d}x = (t+1)\left[\frac{1}{\left(\frac{\delta}{\eta}+2\right)^t} - \frac{1}{\left(\frac{\delta}{\eta}+\frac{R-\delta}{\eta}+2\right)^t}\right]$$

$$= (t+1)\frac{1}{\sqrt[N]{n^t}}\left[\frac{\left(1+\frac{2}{\sqrt[N]{n}}\right)^t - \left(\frac{\delta}{R}+\frac{2}{\sqrt[N]{n}}\right)^t}{\left(1+\frac{2}{\sqrt[N]{n}}\right)^t\left(\frac{\delta}{R}+\frac{2}{\sqrt[N]{n}}\right)^t}\right]$$

using simplification to obtain the last expression. For $n \to \infty$ this results in

$$(t+1)\frac{1}{\sqrt[N]{n^t}}\left(\left(\frac{R}{\delta}\right)^t - 1\right). \tag{33}$$

In the same way we get

$$\sum_{i=0}^{i_{\max}^W} \frac{1}{(\alpha_n^W+i)^{t+1}} \geqslant (t+1)\frac{1}{\sqrt[N]{n^t}}\frac{(\delta/R)^t-(\delta_0/R)^t}{(\delta\delta_0/R^2)^t}. \tag{34}$$

Finally, computing the fraction (32) by (33) and (34) leads to

$$\frac{\delta_0^t(R^t-\delta^t)}{\delta^t(R^t-\delta_0^t)} < \left(\frac{\delta_0}{\delta}\right)^t.$$

This expression can be less than arbitrary $\varepsilon > 0$ using similar arguments as for (28), which completes the proof. □

Remark 5. As we mentioned in Section 2, in both theorems we used the stabilized version (5) of the general KH interpolation (4). When $N = 1$ (16) yields (4) with some apparent substitutions: x_k for A_i, $f(x_k)$ for $\inf B_i$, $\sup B_i$, x for A^* and $|x - x_k|$ for $d_{\mathrm{L}}(A^*, A_i)$, $d_{\mathrm{U}}(A^*, A_i)$.

Remark 6. The KH interpolation function is a rational fraction with order N over N. Hence, its computational complexity is not significantly higher than those of the classical interpolation algorithms.

Remark 7. Theorem 4 is sharp in the sense that for exponents less than N in the expression of distance, the fraction $V/(V+W)$ converges to a constant, as it will be shown in the next example.

The stability result of the stabilized KH interpolation is the generalization of Shepard interpolation, because the stability is valid for the fuzzy-to-fuzzy function approximated by the stabilized KH interpolation, or in other words, for a family of real-valued functions at each $\alpha \in (0,1]$. Moreover, the stability theorems 3 and 4 are established for uniformly distributed node points, which relax the condition of the equispaced node point system for Shepard interpolation. In our future work, we expect to determine the optimal order of convergence in the approximation of the fuzzy-to-fuzzy function exploiting the related results for Shepard interpolation.

It is worth noting that the stability of stabilized KH controllers can be viewed from the aspect of universal approximation capability. Without going deeply into the details, we can state that Theorems 3 and 4 imply that the stabilized KH controllers are capable of approximating any continuous function on a compact domain with respect to the supremum or L_p ($p \in [1,\infty)$) norm. This statement contributes to the topic of the universal approximation capabilities of fuzzy controllers (see e.g. [7,19,32]). It shows that even relaxing the condition of full α-covering property of a rule base, fuzzy systems in general preserve their approximation capabilities.

Now, we turn to show that the theorems are sensitive to the value of distance's quotient. Let us consider the bivariate function $f(x,y)=(x^2+y^2)^c$, where $c>0$, on the unit square. Let us choose equidistant distribution of the measurement points $x_k^{(n)}$ (see Fig. 1). We will show that the approximating function K_n^λ in (29) does not converge to the approximated function f when $\lambda<N$. For simplicity, we approximate the function f in the point $(0,0)$. Hence substituting this value in (18)

$$K_n^\lambda(f,(0,0)) - f(0,0) = \sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\left(\left(\frac{i}{n}\right)^2+\left(\frac{j}{n}\right)^2\right)^c \frac{\dfrac{1}{\left(\left(\frac{i}{n}\right)^2+\left(\frac{j}{n}\right)^2\right)^{\lambda/2}}}{\left(\displaystyle\sum_{i,j=1}^{n-1}\frac{1}{\left(\left(\frac{i}{n}\right)^2+\left(\frac{j}{n}\right)^2\right)^{\lambda/2}}\right)}$$

$$= \frac{\displaystyle\sum_{i,j=1}^{n-1}\left(\left(\frac{i}{n}\right)^2+\left(\frac{j}{n}\right)^2\right)^{c-\lambda/2}}{\displaystyle\sum_{i,j=1}^{n-1}\frac{1}{\left(\left(\frac{i}{n}\right)^2+\left(\frac{j}{n}\right)^2\right)^{\lambda/2}}} = \frac{\displaystyle\sum_{i,j=1}^{n-1}\left(\left(\frac{i}{n}\right)^2+\left(\frac{j}{n}\right)^2\right)^{c-\lambda/2}}{n^\lambda\displaystyle\sum_{i,j=1}^{n-1}\frac{1}{(i^2+j^2)^{\lambda/2}}}. \tag{35}$$

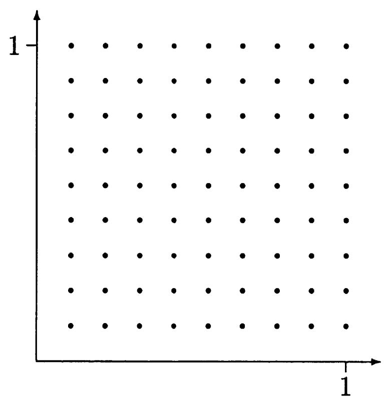

Fig. 1. The distribution of measurement points.

Let us investigate the case when $\lambda = 1 < N(=2)$ and $c = 1/2$. Thus (35) becomes

$$= \frac{\sum_{i,j=1}^{n-1}\left(\left(\frac{i}{n}\right)^2+\left(\frac{j}{n}\right)^2\right)^0}{n\sum_{i,j=1}^{n-1}\frac{1}{(i^2+j^2)^{1/2}}} = \frac{(n-1)^2}{n\sum_{i,j=1}^{n-1}\frac{1}{(i^2+j^2)^{1/2}}} < \frac{1}{\frac{1}{n}\sum_{i,j=1}^{n-1}\frac{1}{(i^2+j^2)^{1/2}}}.$$

Let us further estimate the denominator

$$\frac{1}{n}\sum_{i,j=1}^{n-1}\frac{1}{(i^2+j^2)^{1/2}} = \sum_{i,j=1}^{n-1}\frac{1}{n^2}\frac{1}{\left(\left(\frac{i}{n}\right)^2+\left(\frac{j}{n}\right)^2\right)^{1/2}} \xrightarrow{n\to\infty} \iint_{[0,1]^2}\frac{1}{\sqrt{x^2+y^2}}\,\mathrm{d}x\,\mathrm{d}y$$

using the fact that $(1/n)^2 \approx \Delta x_i \Delta y_i$ is the distance between the measurement points. Finally, we get

$$0 < \iint_{[0,1]^2}\frac{1}{\sqrt{x^2+y^2}}\,\mathrm{d}x\,\mathrm{d}y \leqslant 2\pi\int_{r=0}^{\sqrt{2}}\frac{r}{4}\frac{1}{r}\,\mathrm{d}r = \frac{\pi}{2}\sqrt{2} = \text{const.},$$

thus the difference in (35) converges to a positive constant.

6. Conclusions

The results presented in Theorems 3 and 4 prove the stability of the input–output function of the interpolative KH controllers. This stability is inherited from the α-cuts to the whole fuzzy control algorithm and so, provides a means for constructing tunable non-linear controllers, obviously a reason for the success of fuzzy controllers. Hopefully, similar results can be established for wider classes of fuzzy controllers.

The stability of the KH controller can be placed in a wider context as has been shown in Sections 3 and 4. As it was pointed out in the paper, the well-known interpolation methods have usually searched for an interpolation function in polynomial (or trigonometrical) form. If one neglects this restriction and takes simple rational functions into consideration a new class of interpolation methods can be obtained providing stability independently from the choice of measurement points. We have introduced a class of function derived from the input–output function of the KH interpolation method that fulfils this property. We have pointed out the relevance of this result concerning the universal approximation capabilities of fuzzy systems, and the relationship between KH and Shepard interpolation (the former being the generalization of the latter).

References

[1] J. Balázs, unpublished manuscript.
[2] P. Baranyi, Fuzzy information reduction methods in control algorithms, Ph.D. Thesis, Technical University of Budapest, 1999 (in Hungarian).
[3] R. Barnhill, Representation and approximation surfaces, in: J. Rice (Ed.), Mathematical Software III, Academic Press, New York, 1977, pp. 69–120.
[4] R. Barnhill, R. Dube, F. Little, Properties of Shepard's surfaces, Rocky Mountain J. Math. 13 (1991) 365–382.
[5] L. Bos, K. Salkauskas, Moving least-squares are Backus–Gilbert optimal, J. Approx. Theory 59 (1989) 267–275.
[6] D. Burkhard, P. Bonissone, Automated fuzzy knowledge base generation and tuning, Proc. IEEE Internat. Conf. on Fuzzy Systems, San Diego, CA, 1992, pp. 179–196.
[7] J.L. Castro, Fuzzy logic controllers are universal approximators, IEEE Trans. SMC 25 (1995) 629–635.
[8] G. Criscuolo, G. Mastroianni, Estimates of Shepard interpolatory procedure, Acta Math. Hung. 61 (1993) 79–91.

[9] B. Della Vecchia, Direct and converse results by rational operators, Constr. Approx. 12 (1996) 271–285.

[10] B. Della Vecchia, G. Mastroianni, Pointwise simultaneous approximation by rational operators, J. Approx. Theory 65 (1991) 140–150.

[11] B. Della Vecchia, G. Mastroianni, J. Szabados, Balázs–Shepard operators on infinite intervals, Ann. Uni. Sci. Budapest, Sectio Comput. 16 (1996) 93–102.

[12] B. Della Vecchia, G. Mastroianni, V. Totik, Saturation of the Shepard operators, Approx. Theory Appl. 6 (4) (1990) 76–84.

[13] D. Dubois, H. Prade, Gradual rules in approximate reasoning, Inform. Sci. 61 (1992) 103–122.

[14] R. Franke, Locally determined smooth interpolation at irregularly space points in several variables, Technical Report, Naval Postgraduate School, Monterey, CA, 1975.

[15] L.T. Kóczy, K. Hirota, Approximate reasoning by linear rule interpolation and general approximation, Internat. J. Approx. Reason. 9 (1993) 197–225.

[16] L.T. Kóczy, K. Hirota, Ordering, distance and closeness of fuzzy sets, Fuzzy Sets and Systems 60 (1993) 281–293.

[17] L.T. Kóczy, K. Hirota, Size reduction by interpolation in fuzzy rule bases, IEEE Trans. SMC 27 (1997) 14–25.

[18] L.T. Kóczy, Sz. Kovács, On the preservation of the convexity and piecewise linearity in linear fuzzy rule interpolation, Technical Report TR 93-94/402, LIFE Chair of Fuzzy Theory, Tokyo Institute of Technology, Yokohama, 1993.

[19] B. Kosko, Fuzzy systems as universal approximators, Proc. IEEE Internat. Conf. on Fuzzy Systems, San Diego, CA, 1992, pp. 1153–1162.

[20] D. McLain, Drawing contours from arbitrary data points, Comput. J. 17 (1974) 318–324.

[21] I. Natanson, Constructive theory of functions, Gostekhizdat, Moscow, 1949 (in Russian); Konstruktív függvénytan, Akadémiai Kiadó, Budapest, 1952 (in Hungarian).

[22] D. Newman, T. Rivlin, Optimal universally stable interpolation, Research Report, IBM, Research Div., 1982.

[23] L. Schumaker, Fitting surface to scattered data, in: G. Lorentz, C. Chui, L. Schumaker (Eds.), Approximation Theory II, Academic Press, New York, 1976, pp. 203–268.

[24] D. Shepard, A two dimensional interpolation function for irregularly spaced data, Proc. 23rd ACM Internat. Conf., 1968, pp. 517–524.

[25] Y. Shi, M. Mizumoto, On Kóczy's interpolative reasoning method in sparse rule bases, Proc. 10th Fuzzy Systems Symp., Osaka, 1994, pp. 211–224.

[26] Y. Shi, M. Mizumoto, Some considerations on Kóczy's interpolative reasoning method, Proc. FUZZ-IEEE/IFES '95, Yokohama, 1995, pp. 2117–2122.

[27] Y. Shi, M. Mizumoto, Some considerations on Kóczy's linear interpolative reasoning method, J. SOFT 8 (1996) 147–157 (in Japanese, with English abstract).

[28] G. Somorjai, On a saturation problem, Acta Math. Acad. Sci. Hungar. 32 (1978) 377–381.

[29] J. Szabados, On a problem of R. DeVore, Acta Math. Acad. Sci. Hungar. 27 (1976) 219–223.

[30] J. Szabados, Direct and converse approximation theorems for Shepard-operator, J. Approx. Theory Appl. 7 (1991) 63–76.

[31] D. Tikk, P. Baranyi, Comprehensive analysis of a new fuzzy rule interpolation method, IEEE Trans. on Fuzzy Systems 8 (3) (2000) 281–296.

[32] L.X. Wang, Fuzzy systems are universal approximators, Proc. IEEE Internat. Conf. on Fuzzy Systems, San Diego, CA, 1992, pp. 1163–1169.