General KH controllers are radial basis function interpolators

Domonkos Tikk[†,‡], László T. Kóczy[†,‡] and Tamás D. Gedeon[‡]

†Dept. of Telecom. & Telematics
Budapest Univ. of Technology and Economics
H-1117 Budapest, Pázmány sétány 1/d, Hungary
{tikk,koczy}@ttt.bme.hu
‡School of Information Technology
Murdoch University
South Street, Murdoch, 6150 W.A., Australia
{tgedeon,dtikk}@murdoch.edu.au

Abstract

This paper connects two thoroughly investigated universal approximator techniques to each other. Recently, it has been shown that the input-output function of the general fuzzy KH interpolation method [1, 2] as well as its modification [3] are stable in the mathematical sense, or in other words, they can be considered as universal approximators with respect to the L_p ($p \in [1, \infty]$) norm in the space of continuous functions, defined on a compact domain under certain very general conditions. Our result points out that general KH controllers can be placed into the context of radial basis function (RBF) approximation. However, due to some previously achieved results for KH controllers, some conditions concerning the radial basis functions in the approximation scheme can be relaxed. The advantages of the use of combined neuro-fuzzy applications are also addressed with special emphasis on the benefit of applying fuzzy rule interpolation approaches.

1 Introduction

The classical approaches of fuzzy control deal with dense rule bases where the universe of discourse is α-covered (at any point x from the input set, there exists a fuzzy set A, for which $A(x) \geq \alpha$ holds, usually $\alpha \geq 0.5$) by the antecedent fuzzy sets of the rule base in each dimension. By tuning the rule base or by having only partial information about the system modelled one can obtain rule bases containing "gaps" in the input space, i.e. in these regions there is no or not sufficient information provided by the if-then rules. In such a sparse rule bases no consequent can be constructed by means of the classical methods (CRI, Mamdani, Larsen, Takagi–Sugeno) if the observation is within a gap. The fuzzy rule interpolation method (first proposed by Kóczy and Hirota, [4]) provides a tool to construct fuzzy controllers handling sparse rule bases. The KH interpolation methods have been investigated thoroughly. Recently, it has been shown [1, 2] that the stabilized general KH interpolation possesses the universal approximation property.

A popular function approximation technique in the field of neural networks is the radial basis function (RBF) network ([5], [6], Chap. 5.). These types of three-layered feedforward networks are also considered to be universal approximators.

In this work we connect these soft computing techniques showing that general KH interpolations have strong connections with the RBF approximation scheme.

The paper is organized as follows. Section 2 recalls the KH interpolation methods and the theorem of its universal approximation property. Section 3 gives an overview of the RBF interpolation and approximation scheme, as well as introduces the approximation capabilities of RBF networks. Section 4 contains the main results of the paper, and finally, Section 5 mentions some conclusions.

2 The KH interpolation methods

Fuzzy rule interpolation was proposed by Kóczy and Hirota (see [4]). This technique was based on the α-cut distances of convex and normal fuzzy sets. The method generates the fuzzy conclusion by means of its α-cuts based on the Extension and the Resolution Principles.

The simplest version of this method is called linear fuzzy rule interpolation of two rules (in each dimension) for the area between their antecedents. This can be applied if the observation is located in between the two rule antecedents with respect to the ordering defined on the input universe. For every important cut, e.g. in the case of triangular and trapezoidal shaped membership functions $\alpha = 0$ and 1 (see [7]), it determines the conclusion

Research supported by the Australian Research Council large grants scheme, by the National Scientific Research Foundation (OTKA) Grants No. T019671, T030655, and by the Hungarian Ministry of Culture and Education (MKM) Grants No. FKFP 0235/1997 and 0422/1997.

as

$$B^*_{\alpha,C} = \frac{\frac{B_{1\alpha,C}}{d_C(A^*_\alpha,A_{1\alpha})} + \frac{B_{2\alpha,C}}{d_C(A^*_\alpha,A_{2\alpha})}}{\frac{1}{d_C(A^*_\alpha,A_{1\alpha})} + \frac{1}{d_C(A^*_\alpha,A_{2\alpha})}}, \quad C \in \{L,U\} \tag{1}$$

where A_i and B_i $(i = 1,2)$ denote antecedents neighbouring the observation A^* and corresponding consequents, respectively, subscripts L and U in C refer to the minimum and maximum of respective α-cuts, and $d_C : \mathbb{R} \times \mathbb{R} \to [0,+\infty)$, $C \in \{L,U\}$ is an appropriate distance function (cf. [8]). This approach is referred to as linear fuzzy KH interpolation.

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. Thus, (1) can be generalized as

$$B^*_{\alpha,C} = \frac{\sum_{i=1}^{m} \frac{1}{d_C(A^*_\alpha,A_{i\alpha})} B_{i\alpha,C}}{\sum_{i=1}^{m} \frac{1}{d_C(A^*_\alpha,A_{i\alpha})}}, \qquad C \in \{L,U\}. \tag{2}$$

More details on this method can be found in [4].

We recall the theorem on the approximation capabilities of general KH interpolation. (For the sake of clarity, henceforth we denote vectors and matrices with bold typesetting.)

Theorem 1 *Consider the L_p norm $\|\cdot\|_p$ with $p \in [1,\infty]$, a positive real $\varepsilon > 0$, the n dimensional 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 for all $\mathbf{x} \in \Omega$*

$$\left| \sum_{i=1}^{m} f(\boldsymbol{\xi}_i) \frac{\frac{1}{\|\mathbf{x}-\boldsymbol{\xi}_i\|_p^n}}{\sum_{j=1}^{m} \frac{1}{\|\mathbf{x}-\boldsymbol{\xi}_j\|_p^n}} - f \right| < \varepsilon \tag{3}$$

where the measurement points $\boldsymbol{\xi}_i$ $(i = 1,\dots,m)$ are uniformly distributed on Ω.

In this theorem we modified the denotation of equations (2) according to the ones in approximation theory. Moreover, instead of taking simply the distance d we consider its nth power d^n, where n denotes the dimension of the domain Ω. This modification of the general KH interpolation is called the stabilized general KH interpolation. When $n = 1$ (3) yields (2) with some apparent substitutions: $\boldsymbol{\xi}_i$ for A_i, $f(\boldsymbol{\xi}_i)$ for $B_{i\alpha,C}$, $\mathbf{x}$ for A^* and $|\mathbf{x}-\boldsymbol{\xi}_i|$ for $d_C(A^*,A_{i\alpha})$ $(C \in \{L,U\})$.

In other words, regarding the stability of (2), this result can be interpreted in the fuzzy context in the following way: If the antecedents or the observation of a rule base changes only slightly it neither does not cause a significant change in the conclusion.

3 Approximation by RBF

The problem of exact real multivariate interpolation can be formulated as follows [5]: given a set of distinct vectors $\{\mathbf{x}_i; i = 1,\dots,m\}$ and m real numbers $\{y_i; i = 1,\dots,m\}$, determine a function fulfilling the condition

$$f(\mathbf{x}_i) = y_i, \qquad i = 1,\dots,m. \tag{4}$$

One of the possible solutions of the equation (4) is the RBF approach. This introduces the so-called radial basis functions, $\phi_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1,\dots,m$ and uses the form

$$\phi_i(\mathbf{x}) = \phi(\|\mathbf{x}-\boldsymbol{\xi}_i\|),$$

($\|\cdot\|$ denotes a norm, usually the Euclidean, on $\mathbb{R}^n$). Here $\phi : [0,\infty) \to \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$ and $\boldsymbol{\xi}_i \in \mathbb{R}^n$ are called the centres of functions ϕ. Functions ϕ_i $(i = 1,\dots,m)$ form a basis of a linear function space, so the interpolating function f can be expressed as their linear combination:

$$f(\mathbf{x}) = \sum_{i=1}^{m} \lambda_i \phi_i(\|\mathbf{x}-\boldsymbol{\xi}_i\|). \tag{5}$$

Substituting equation (5) into (4), a set of linear equations for the coefficients λ_i are obtained [5]:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} \phi_{11} & \dots & \phi_{1m} \\ \vdots & \ddots & \vdots \\ \phi_{m1} & \dots & \phi_{mm} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \cdots \\ \lambda_m \end{pmatrix} \tag{6}$$

By denoting the matrices in the equation (6) by $\mathbf{y}$, $\boldsymbol{\Phi}$, and $\boldsymbol{\lambda}$, respectively, this can be simply written as

$$\mathbf{y} = \boldsymbol{\Phi}\boldsymbol{\lambda},$$

where

$$\boldsymbol{\Phi} = (\phi_{ij}) = \phi(\|\mathbf{x}_i-\boldsymbol{\xi}_j\|), \qquad i,j = 1,\dots,m.$$

From here the unique solution of (6) can be determined as

$$\boldsymbol{\lambda} = \boldsymbol{\Phi}^{-1}\mathbf{y}. \tag{7}$$

if matrix $\boldsymbol{\Phi}$ is nonsingular (i.e. regular).

In general, the regularity of matrix $\boldsymbol{\Phi}$ can not be ensured for arbitrary ϕ. Nevertheless, according to the results of Micchelli and Powell ([9], [10]) a large class of functions ensures the regularity of matrix $\boldsymbol{\Phi}$ under the condition of distinct $\boldsymbol{\xi}_i$. This class includes, e.g. the following functions:

1. $\phi(z) = z$;
2. $\phi(z) = z^2 \log z$;

3. $\phi(z) = z^3$;
4. $\phi(z) = (z^2 + c)^\alpha$, $c \geq 0$, $0 < \alpha < 1$;
5. $\phi(z) = \exp(-z^2/c)$, $c \geq 0$;
6. $\phi(z) = (z^2 + c)^{-1/2}$, $c \geq 0$.

These functions belong to two different groups: *localized* functions which satisfy $\lim_{z \to 0} \phi(z) = 0$, and so $\mathbf{\Phi}$ is a positive definite; while *non-local* functions are unbounded when $z \to 0$, and hence, $\mathbf{\Phi}$ is not positive definite, having $m - 1$ negative eigenvalues and only one positive one [9].

The RBF approach can be easily extended to multiple output mappings.

The interpolation scheme just outlined has the disadvantage that the number of RBFs ϕ is identical with the number of data points. In the case where the number of data points is significantly larger, we have an approximation problem.

In this context the inequality $n_0 \ll m$ holds for the number of data points m, and for n_0, the number of RBFs ϕ. In this case the problem becomes overspecified. In this case the function ϕ_i are calculated by using the Moore–Pensore pseudo inverse ([5, 6])

$$\mathbf{\Phi}^+ = (\mathbf{\Phi}^T \mathbf{\Phi})^{-1} \mathbf{\Phi}^T,$$

where $\mathbf{\Phi}^T \mathbf{\Phi} = \mathbf{I}_m$, the $m \times m$ identity matrix. We note, that in this case the centres $\boldsymbol{\xi}_i$ can not be defined straightforwardly as a certain set of data points.

The approximation problem of RBFs has been investigated thoroughly in approximation and neural networks theory. Here we briefly outline the approximation scheme implemented by feedforward neural networks with RBF units [5]. For further details, we refer the Reader to the comprehensive overview in [6], Chap. 5.

The RBF approximation problem can be naturally represented by a three-layer feedforward network with fully interconnected layers. The n input units correspond to the n-dimensional input vectors. The connections between the input and hidden units are weighted by the appropriate coordinate of the corresponding centre. The hidden units calculate the functions ϕ in (5). The links between the hidden an the output nodes are weighted by the coefficients λ. The output units are linear calculating the weighted sum of their inputs in accordance with (5). Concerning the training algorithms of RBF networks we refer to [6].

The universal approximation property of certain RBF networks has been proved by several authors [11, 12, 13]. These results differs slightly regarding the applied conditions for the radial function ϕ.

4 Main result

Now we describe how the previous two techniques can be connected.

Proposition 1 *The input-output function* (3) *of the stabilized general KH interpolator determines an RBF interpolation scheme* (expression (5)) *with the radial functions $\phi_{\mathrm{sgKH}}(z) = \frac{1}{z^n}$ and the following coefficients*

$$\lambda_i = \frac{y_i}{\sum_{j=1}^{m} \frac{1}{\|\mathbf{x} - \boldsymbol{\xi}_j\|^n}} = \frac{y_i}{\sum_{j=1}^{m} \phi_{\mathrm{sgKH}}(\|\mathbf{x} - \boldsymbol{\xi}_j\|)}. \tag{8}$$

Proof. It can be obtained immediately substituting the coefficients (8) into the formula (5) with radial functions ϕ_{sgKH}. So, in this case, the unique solution of the linear equation system (6) is given explicitly by (8). Moreover, as a consequence of the approximation theorem for KH interpolators, the vectors $\boldsymbol{\xi}_j$ $(j = 1, \ldots, m)$ need not to be distinct, but uniformly distributed on compact subset of $\mathbb{R}^n$.

As we did not exploit here the result of Theorem 1, similar constructions apply for the simpler general KH interpolation.

Proposition 2 *The input-output function* (2) *of the general KH interpolator determines a RBF interpolation scheme* (expression (5)) *with the radial functions $\phi_{\mathrm{gKH}}(z) = \frac{1}{z}$ and the following coefficients*

$$\lambda_i = \frac{y_i}{\sum_{j=1}^{m} \frac{1}{\|\mathbf{x} - \boldsymbol{\xi}_j\|}} = \frac{y_i}{\sum_{j=1}^{m} \phi_{\mathrm{gKH}}(\|\mathbf{x} - \boldsymbol{\xi}_j\|)}. \tag{9}$$

According to these result, general KH controllers can be implemented as RBF networks. With a neuro-fuzzy method, we can exploit the advantageous properties of both neural and fuzzy techniques.

In the case where one has certain knowledge about the modelled system, e.g. interpreted as fuzzy rules, this information can be used to determine the initial parameters of a neural network. By means of the training data and the learning algorithm applied to the network, the corresponding fuzzy rule base can be tuned, this way, refining the model of the system. Such an approach results in a well-tuned model of the system (naturally to some extent depending on the reliability and appropriateness of the training data), while the tuned rule base model of the system can still provide a linguistic description of the system, thus, it assures the tracing of the system by a human operator. The linguistic description of the system gives a possible tool for quality control in certain cases: if the training data set is noisy or degenerated, a human operator can discover this kind of mistakes much easier based on a linguistic description than merely on the transfer function calculated by the network.

In the case when only input-output sample data are available, a combined neuro-fuzzy approach helps in the extraction of fuzzy rules from the data. Once the rules are extracted from the training data set the rule based description of the system is available.

The advantage of the use of a fuzzy rule interpolation approach compared to the classical fuzzy reasoning approach is that one can have such input-output sample data sets which does not cover the whole range of the input space, e.g. in large regions of the input space the appropriate output remains unknown. As a consequence, the extracted rule base can contain gaps (a sparse rule base). In such a case the fuzzy controllers operating on dense rule bases can not be applied. As a practical example we refer to the well-log analysis in petroleum industry (see [14]).

5 Conclusion

In this paper a new example has been given for the approximate functional equivalence of fuzzy and neural models. We have shown that two techniques, the general fuzzy KH interpolation and the RBF approximation scheme (usually implemented by RBF networks); both being considered as universal approximators, can be connected. The general fuzzy KH interpolation can be considered as a special RBF approximation scheme for fixed α-cuts and $C \in \{L, U\}$. The possible benefits of implementing a fuzzy rule interpolation technique by a neural network is also outlined.

Acknowledgement

The authors want to thank to János Levendovszky for his valuable suggestions on the connection of KH interpolation and RBF functions.

References

[1] D. Tikk, I. Joó, L. T. Kóczy, P. Várlaki, B. Moser, and T. D. Gedeon. Stability of interpolative fuzzy KH-controllers. To appear in *Fuzzy Sets and Systems*.

[2] D. Tikk, L. T. Kóczy, and B. Moser. Stability of interpolative fuzzy KH-controllers. Multi-dimensional case. In *Proc. of the Tempus JEP Modify Symposium*, page 8, Budapest and Miskolc, 1997.

[3] D. Tikk, P. Baranyi, Y. Yam, and L. T. Kóczy. Stability of a new interpolation method. In *Proc. of the IEEE Int. Conf. on System, Man, and Cybernetics (IEEE SMC'99)*, volume III, pages 7–9, Tokyo, Japan, 1999.

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

[5] K. Hlavačková and R. Neruda. Radial basis functions. Technical report, Institute of Information and Computer Science, Czech Academy of Science, Prague, w.y.

[6] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, Upper Saddle River, New Jersey, 1999.

[7] L. T. Kóczy and Sz. Kovács. The convexity and piecewise linearity of the fuzzy conclusion generated by linear fuzzy rule interpolation. *BUSEFAL*, 62(Automne):23–29, 1994.

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

[9] C. A. Micchelli. Interpolation of scattered data: Distance matrices and conditionally positive definite functions. *Construct. Approx.*, 2:11–22, 1986.

[10] M. J. D. Powell. The theory of radial basis function approximation in 1990. In W. Light, editor, *Advances in Numerical Analysis Vol. II: Wavelets, Subdivision Algorithms, and Radial Basis Functions*, pages 105–210. Oxford Science Publications, Oxford, 1992.

[11] E. J. Hartman, J. D. Keeler, and J. M. Kowalski. Layered neural networks with Gaussian hidden units as universal approximations. *Neural Computation*, 2:210–215, 1990.

[12] J. Park and I. W. Sandberg. Universal approximation using radial-basis-function networks. *Neural Computation*, 3:246–257, 1991.

[13] T. Poggio and F. Girosi. Networks and the best approximation property. *Biological Cybernetics*, 63:169–176, 1990.

[14] K. W. Wong and T. D. Gedeon. Petrophysical properties prediction using self-generating fuzzy rules inference system with modified α-cut based fuzzy interpolation. To Appear in the *Proc. of ICONIP 2000* to be held in Taejon, Korea, 2000.