Neurocomputing 149 (2015) 265–274

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Approximation properties of ELM-fuzzy systems for smooth functions and their derivatives ☆

De-Gang Wang [a,b,*], Wen-Yan Song [c], Hong-Xing Li [a]

[a] School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China
[b] Informedia Electronic Co., Ltd. Dalian 116001, China
[c] School of Mathematics and Quantitative Economics, Dongbei University of Finance and Economics, Dalian 116025, China

ARTICLE INFO

Article history:
Received 2 August 2013
Received in revised form
28 January 2014
Accepted 4 February 2014
Available online 10 September 2014

Keywords:
Fuzzy systems
Generalized Bernstein polynomials
Spline function
Extreme learning machine
Universal approximation

ABSTRACT

In this paper, we utilize generalized Bernstein polynomials to construct fuzzy system. Different from traditional Bernstein polynomials, partition of interval on input variable can be chosen as non-equidistant division. We prove that generalized Bernstein fuzzy systems are universal approximators to a given continuous function and its high-order derivatives. Further, ELM method is used to tune the parameters of generalized Bernstein fuzzy system and Spline fuzzy system. It is proved that ELM-Spline fuzzy system can approximate a function and its derivative. Simulation examples show that the proposed ELM-Bernstein fuzzy system and ELM-spline fuzzy system can achieve high approximation capability for nonlinear model.

1. Introduction

Fuzzy system theory is a useful tool to deal with some complex systems with fuzziness and linguistic variables. In practical application, one of the main design objectives is to construct fuzzy system to approximate a desired model.

As fuzzy system can be designed by both linguistic information and input–output data, in recent years, universal approximation theory of fuzzy system has gained wide attention from scholars and engineers, and it has been applied in many fields, for example, in control systems [1,2], in model approximation and prediction [3,4] and in decision making [5]. In order to demonstrate the universal approximation capabilities of fuzzy systems, many scholars attempt to construct various fuzzy systems to approximate nonlinear function or dynamic model. In [6,7], it is proved that Mamdani fuzzy systems are universal approximators to nonlinear functions. And, in order to measure the approximation capability of Mamdani fuzzy system quantitatively, the approximation error bounds of fuzzy systems are established in [8,9]. Further, to solve the rule explosion problem of fuzzy system, hierarchical fuzzy systems are designed to approximate function in [10–12]. Different from Mamdani fuzzy systems, the consequents of TS fuzzy systems are represented as the functions of input variables. In [13], TS fuzzy systems with linear consequent are proved to be universal approximators to nonlinear function. In [14,15], formulae which can determine the minimal upper bounds on the number of fuzzy rules for TS fuzzy systems are established. Besides, the approximation properties of other types of fuzzy systems, such as Boolean fuzzy systems [16] and generalized Bernstein fuzzy systems [17], are investigated. In real world applications, many models, such as control system and decision model, operate in dynamic environment. Thus, some novel fuzzy systems are developed to approximate various classes of dynamic models. In [18], probabilistic fuzzy logic system is proposed for the dynamic modeling problem in stochastic circumstance. It is pointed out that fuzzy systems can be capable of approximating a class of autonomous systems [19]. Furthermore, in [20], time variant fuzzy system is proposed to approximate non-autonomous system.

Although fuzzy system can describe uncertain and complex system, in some application, it still requires that the proposed model should be universal approximators not only for the given smooth function but also for its derivative. Indeed, a great amount of efforts have been devoted to this topic, and many results have been obtained. Some scholars utilize various membership functions to construct fuzzy systems to achieve the smooth approximation capability. In [21,22], fuzzy systems with Gaussian membership

☆This work was supported by the National Natural Science Foundation of China (71201019, 61104038, 61374118), the China Postdoctoral Science Foundation funded project (2013M541233), the Fundamental Research Funds for the Central Universities (DUT14QY30) and the Program for Liaoning Excellent Talents in University (WJQ2014036).

* Corresponding author at: School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China.
E-mail address: wangdg@dlut.edu.cn (D.-G. Wang).

http://dx.doi.org/10.1016/j.neucom.2014.02.070
0925-2312/

functions and π functions are respectively established to approximate the function and its derivative. And, in [23], the translations and dilations of one fixed function are chosen as the basis functions to construct fuzzy system which possesses smooth approximation property. Further, in [24] fuzzy transform technology is also introduced to approximate the function and its derivative.

Theoretically, neural network can also possess universal approximation capability for nonlinear function in [25–27]. Further, a novel single hidden layer feedforward network (SLFN) with random hidden nodes, called Extreme Learning Machine (ELM), is proposed in [28]. Different from the traditional neural network and traditional feedforward network learning algorithms, the hidden layer parameters in ELM [29] need not be learned and the smallest train error can be achieved. In [29–31], it is proved that ELM possesses interpolation capability and universal approximation capability. In addition, the smooth approximation properties of neural network have also attracted attentions by many scholars. In [32], it is proved that multilayer feedforward networks with smooth activation functions can approximate the function and its derivative. Further, in order to improve the approximation accuracy, in [33] nonsigmoid activation functions are utilized to construct the multilayer feedforward neural network. In [34], a probabilistic method based on central limit theorems is applied to demonstrate the smooth approximation capability of neural network. In [35], simultaneous approximation of multivariate function and its partial derivatives is investigated. In addition, in [36], algebraic constructive approach is utilized to achieve the simultaneous approximation of neural network, and the approximation error bound of the corresponding neural network is established.

The above works show that both fuzzy system and neural network possess the capabilities of simultaneous approximations for functions and their derivative. However, how to construct a hybrid system which combines fuzzy system with neural network to provide simultaneous approximation of function and its higher derivatives is still an interesting and worthy of discussion question. In order to solve this question, we need to choose appropriate basis function and learning algorithm. It is well known that in numerical analysis, Bernstein polynomial and spline function are widely used in function approximation and differential equation solving. Motivated by the above facts, in this paper, we choose generalized Bernstein polynomial and spline function to design fuzzy system and apply ELM to train the corresponding parameters. Compared to the traditional Bernstein polynomials and spline function, the proposed ELM-fuzzy system can obtain higher smooth approximation accuracy. Meanwhile, for some commonly used basis functions, such as radial basis function (RBF) and sigmoid function, ELM-fuzzy systems with generalized Bernstein polynomial or spline basis function can still possess a better approximation capability.

The paper is organized as follows. In Section 2, some preliminaries are introduced. In Section 3, the approximation properties of generalized Bernstein fuzzy systems are investigated. In Section 4, ELM-fuzzy systems are constructed to approximate functions and their derivatives. Further, in Section 5, some numerical examples are carried out to illustrate the validity of the results. In Section 6, some concluding remarks are then provided.

2. Preliminaries

In this section, we introduce several notations and concepts which are used in this paper.

The space of continuous function on a closed bounded interval $[a,b]$ is defined by $C[a,b]$. If n is a positive integer, then the space of n-times continuously differentiable functions on $[a,b]$ is defined by $C^n[a,b]$, particularly $C^0[a,b]=C[a,b]$.

Definition 2.1 (*Mo and Liu [37]*). Suppose that $f \in C[0,1]$, then the Bernstein polynomials of degree $n (n \geq 1)$ is defined as

$$(B_n f)(x) = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right). \tag{1}$$

In [37], B_n is also called as the Bernstein operator.

Definition 2.2 (*Mo and Liu [17]*). Suppose that $f \in C[0,1]$, then the generalized Bernstein polynomial of degree $n (n \geq 1)$ is defined as

$$(\overline{B}_n f)(x) = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k} f(x_k),$$

where $0 = x_0 < x_1 < \cdots < x_n = 1$.

Different from Bernstein polynomials (1), the partition points $x_k (k=0,\ldots,n)$ of the generalized Bernstein polynomials may not be equidistant partition, which means that Bernstein polynomials are special cases of generalized Bernstein polynomials.

Definition 2.3 (*Mo and Liu [38]*). Consider an interval $[a,b]$, and subdivide it by a mesh of points corresponding to the locations of the ducks $\Delta : a = x_0 < x_1 < \cdots < x_n = b$. A function $S_\Delta(x)$ is said to be a polynomial spline of degree m with respect to the mesh Δ, or a spline on Δ, if $S_\Delta(x)$ is satisfied:

(1) in each subinterval $[x_{k-1}, x_k], (k=1,\ldots,n)$, $S_\Delta(x)$ coincides with a polynomial of degree m;
(2) $S_\Delta^{(j)}(x)\ (j=0,\ldots,m-1)$, which are all the jth derivatives of $S_\Delta(x)$, are all continuous on $[a,b]$.

It is well known that a spline is a sufficiently smooth polynomial function that is piecewise-defined and possesses a high degree of smoothness at the knots where the polynomial pieces connect. In addition, if an associated set of ordinates is prescribed as Y: $y_0, y_1, \ldots, y_n$, and the polynomial spline $S_\Delta(x)$ satisfies that $S_\Delta(x_k) = y_k\ (k=0,\ldots,n)$, then $S_\Delta(x)$ is an interpolating polynomial spline on the mesh Δ.

Example 1. If the mesh $\Delta : a = x_0 < x_1 < \cdots < x_n = b$ on $[a,b]$ is an equidistant partition, which means that $x_k = a + k \cdot h\ (k=0,\ldots,n)$ and $h = (b-a)/n$, then the typical cubic uniform spline functions can be defined as

$$\Omega_k(x) = \begin{cases} \frac{1}{2}\left|\frac{x-x_k}{h}\right|^3 - \left(\frac{x-x_k}{h}\right)^2 + \frac{2}{3}, & \left|\frac{x-x_k}{h}\right| \leq 1, \\ -\frac{1}{6}\left|\frac{x-x_k}{h}\right|^3 + \left(\frac{x-x_k}{h}\right)^2 - 2\left|\frac{x-x_k}{h}\right| + \frac{4}{3}, & 1 < \left|\frac{x-x_k}{h}\right| \leq 2, \\ 0, & \left|\frac{x-x_k}{h}\right| > 2. \end{cases}$$

If the above mesh Δ is expanded as $x_{-1} < x_0 < x_1 < \cdots < x_n < x_{n+1}$, where $x_k = a + k \cdot h\ (k=-1,0,\ldots,n,n+1)$ and $h=(b-a)/n$, then the sequence of cubic spline functions $\{\Omega_k\}\ (k=-1,0,\ldots,n,n+1)$ forms a basis for the space of cubic polynomial spline $S_\Delta(x)$. Besides, for any $x \in [a,b]$, $\Omega_k(x) \geq 0$ and $\sum_{k=-1}^{n+1} \Omega_k(x) = 1$.

In applied mathematics, generalized Bernstein polynomials and spline functions are widely applied in many fields, such as numerical approximation and nonlinear differential equation solving. And these two functions are also bounded non-constant continuous functions which satisfy the condition of the activation function in ELM. Hence, in the following, we will choose these two functions to construct ELM fuzzy system.

Then, we take a single-input-single-output system as an example to introduce the construction of fuzzy system. We respectively denote $X=[a,b]$ and $Y=[c,d]$ as the input space and the output