Neurocomputing 190 (2016) 237–242

Contents lists available at ScienceDirect

Neurocomputing

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

Scattered data approximation by neural networks operators

Zhixiang Chen [a], Feilong Cao [b,*]

[a] Department of Mathematics, Shaoxing University, Shaoxing 312000, Zhejiang Province, PR China
[b] Department of Applied Mathematics, China Jiliang University, Hangzhou 310018, Zhejiang Province, PR China

ARTICLE INFO

Article history:
Received 11 August 2015
Received in revised form
3 January 2016
Accepted 14 January 2016
Communicated by D. Wang
Available online 1 February 2016

Keywords:
Neural networks
Approximation
Interpolation
Scattered data
Error

ABSTRACT

In this paper, some feed-forward neural networks (FNNs) interpolation operators based on scattered data are introduced. Further, these operators are used as approximators to approximate bivariate continuous target function. By means of the translations and dilates of logistic function, some FNNs quasi-interpolation and exact interpolation operators are constructed, respectively. Using the modulus of continuity of function and the mesh norm of scattered data as measures, the corresponding approximation errors of the constructed operators are estimated. In addition, the well-known central B-splines are used to construct FNNs interpolation operators with compact support, and the corresponding approximation errors are also estimated.

1. Introduction

In the theory and application of feed-forward neural networks (FNNs), activation function of networks plays an important role, which is often taken as sigmoidal function, i.e., the bounded function on $\mathbb{R}$ satisfying

$$\lim_{x\to-\infty}\sigma(x)=0, \quad \lim_{x\to+\infty}\sigma(x)=1.$$

Among numerous sigmoidal functions, the logistic function defined by

$$\sigma(x)=\frac{1}{1+e^{-x}} \tag{1.1}$$

is typical, and which has been used extensively as activation function of FNNs.

As we know, FNNs with sigmoidal activation function are universal approximator. Theoretically, any continuous or Lebesgue integrable function defined on a compact set can be approximated by an FNN with sigmoidal activation function to any desired degree of accuracy by increasing the number of hidden neurons, which is usually called density problem of FNNs approximation. At present, this problem has been satisfactorily solved, and many classical results can be found in more articles, such as Cybenko [31], Funahashi [32], Hornik et al. [34,35], Hornik [36,37], Ito [59], Chui and Li [22], Kurková [38], Leshno et al. [43], Park and Sandberg [49,50], Chen [13–15], Chen and Chen [16–18], Chen et al. [20,19], and Pinkus [51]. Another important problem for such approximation is called complexity problem, which mainly describes the relationship between the approximation error and the number of neurons in hidden layer. This problem has attracted a lot of interests of researchers. We refer readers to [7,12,39–41,6,47,48,53,57,58,10,21].

To study the complexity problem, the FNNs operators usually are constructed for approximating the target function, and the approximation errors are estimated. In 1992, Cardaliaguet and Euvrard [11] first introduced an interesting neural network operator called Cardaliaguet Euvrard operator by using the centered bell-shaped continuous function with compact support. In 1997, Anastassiou [1] discussed the approximation error of the operators. More researchers, such as Attali and Pagès [6], Suzuki [53], Hahm and Hong [33], Lewicki and Marino [44], Xu and Cao [57,58], and Chen and Cao [21], constructed various FNNs operators with the sigmoidal function and studied their approximation properties. Recently, Costarelli [23,24] and Costarelli and Spigler [25–30] made a series of investigation on the univariate or multivariate FNNs operators with sigmoidal activation functions, and obtained some interesting results. Using a class of specific sigmoidal function called hyperbolic tangent function as activation function of FNNs, Anastassiou [2–5] made some in-depth research on the construction and approximation of FNN operators.

On the other hand, it is well known that interpolation is a class of important method to approximate or fit function all the time. Llanas and Sainz [45] first applied the interpolation idea into FNNs and constructed a class of interpolation FNNs with nondecreasing

* Corresponding author.
E-mail address: feilongcao@gmail.com (F. Cao).

http://dx.doi.org/10.1016/j.neucom.2016.01.013
0925-2312/

activation function that could approximately interpolate univariate or multivariate target functions. Such type of FNNs, in fact, can be regarded as quasi-interpolation operators from mathematical viewpoint. In [23], Costarelli constructed a kind of interpolation FNNs with ramp and central B-splines activation functions to approximate continuous target functions defined on bounded intervals, and in [24], he further studied the case that the activation function is the tensor product of B-splines, and proved that the constructed FNNs operators could interpolate the target functions on uniform or non-uniform spaced grid.

All the researches on FNNs interpolation mentioned above are based on the nodes (samples) of rule distribution. In practical application, however, there are a large number of scattered data. Since one does not have any control on where these data (sampling points) site, compared with the interpolation with nodes (samples) of rule distribution, it is difficult to tackle the interpolation based on scattered data [54–56,60]. The main purpose of this paper is to study the approximation of scattered data by means of FNNs interpolation operators. We will construct some FNNs exact interpolation or quasi-interpolation operators based on scattered data to approximate bivariate function. The activation function of constructed FNNs operators is generated by the translation and dilatation of the logistic function, and thus it is easy to compute and apply in reality. Also, we will utilize some techniques of mathematical analysis and methods of approximation theory to estimate the approximation error of the operators. In addition, we will make some discussions on the construction and approximation of FNNs operators with B-splines activation function.

The rest of this paper is organized as follows. In Section 2, we will study the construction and approximation of FNNs quasi-interpolation operators with activation function generated by the translation and dilatation of the logistic function. In Section 3, we will construct the FNNs exact interpolation operators by the activation functions given in Section 2, and estimate the approximation error. In Section 4, we will use B-splines to construct FNNs quasi-interpolation operators and obtain the estimation of approximation error of the operators. Finally, some conclusions and remarks will be arranged in Section 5.

2. The construction and approximation of FNNs quasi-interpolation operators

For the logistic function given in (1.1), setting

$$g_\sigma(x):=\tfrac{1}{2}(\sigma(x+1)-\sigma(x-1)),$$

then we have [21]

- $g_\sigma(x)=\frac{e-e^{-1}}{2}\cdot\frac{e^x}{(1+e^{x+1})(1+e^{x-1})}$.
- $g_\sigma(x)>0$.
- $g_\sigma(x)$ is an even function.
- $g_\sigma(x)$ is non-decreasing for $x<0$ and non-increasing for $x>0$.
- The support of g_σ, $\operatorname{supp}(g_\sigma)$, is $\mathbb{R}$.

Let $x_i\in[-1,1]^2\subset\mathbb{R}^2$ $(i=1,2,\dots,N)$ be sample points, which constitute the interpolation nodes of samples: $(x_1,f_1),(x_2,f_2),\dots,(x_N,f_N)$. When $x_1,x_2,\dots,x_N$ are scattered, that is, they are not any uniform spaced grid points, we call $(x_1,f_1),(x_2,f_2),\dots,(x_N,f_N)$ scattered data. To characterize scattered data points set $X:=\{x_1,x_2,\dots,x_N\}$, we need three quantities [54]: mesh norm h_X, separation radius q_X, and mesh ratio ρ_X, which are defined as

$$h_X:=\sup_{y\in[-1,1]^2}\inf_{\xi\in X}d(\xi,y),\quad q_X:=\frac{1}{2}\min_{\xi\neq\xi'}d(\xi,\xi'),\quad \rho_X:=\frac{h_X}{q_X},$$

respectively, where $d(\cdot,\cdot)$ denotes the Euclidean distance. Obviously, $\rho_X\ge 1$. If there exists a ρ independent of X such that $\rho_X\le\rho$, then we say that X is ρ-uniform. And in this section and Section 3, we all assume that X is ρ-uniform.

We use $B(x_i,h_X)$ to denote the closed ball with radius h_X and center x_i. It is not difficult to know that

$$[-1,1]^2\subseteq\bigcup_{i=1}^N B(x_i,h_X).$$

We also denote the set of all continuous functions defined on $[-1,1]^2$ by $C_{[-1,1]^2}$, which forms a Banach space with norm

$$\|f\|:=\max_{x\in[-1,1]^2}|f(x)|.$$

The modulus of continuity of $f\in C_{[-1,1]^2}$ is defined as [46]

$$\omega(f,t):=\sup_{\|h\|_2\le t}\ \max_{x+h,x\in[-1,1]^2}|f(x+h)-f(x)|,\quad t>0,$$

where $\|\cdot\|_2$ is the Euclidean norm. The modulus of continuity of continuous function f is an important measure to depict the continuity and smoothness of function f, which plays an important key role in the approximation theory and harmonic analysis, and it is often used as a metric to estimate the approximation error. Specially, if for an α $(0<\alpha\le 1)$ there is a constant $C>0$, such that $\omega(f,t)\le Ct^\alpha$ $(t\to 0^+)$, then we say that f is a Lipschitz function and write $f\in \mathrm{Lip}_C\alpha$.

For the scattered points $x_j\in[-1,1]^2(1\le j\le N)$, the logistic function given in (1.1), and parameter $\lambda>0$, we define

$$g_j^\lambda(x):=\frac{1}{2}\left(\sigma\left(\frac{\|x-x_j\|_2}{\lambda}+1\right)-\sigma\left(\frac{\|x-x_j\|_2}{\lambda}-1\right)\right).\tag{2.1}$$

Then

$$g_j^\lambda(x)=\frac{e-e^{-1}}{2}\cdot\frac{e^{\frac{\|x-x_j\|_2}{\lambda}}}{\left(1+e^{\frac{\|x-x_j\|_2}{\lambda}+1}\right)\left(1+e^{\frac{\|x-x_j\|_2}{\lambda}-1}\right)},\quad j=1,2,\dots,N.\tag{2.2}$$

For any function f defined on $[-1,1]^2$, scattered points $x_j\in[-1,1]^2(1\le j\le N)$, and $g_j^\lambda(x)$ given by (2.1), we define quasi-interpolation operators

$$(Q_N^\lambda f)(x):=\sum_{i=1}^N f(x_i)\frac{g_i^\lambda(x)}{\sum_{j=1}^N g_j^\lambda(x)}.\tag{2.3}$$

Now we estimate the approximation error of quasi-interpolation operators $(Q_N^\lambda f)$ given in (2.3) approximating continuous function f defined on $[-1,1]^2$. From the definition (2.3), we have

$$\begin{aligned}|f(x)-(Q_N^\lambda f)(x)| &\le \sum_{i=1}^N|f(x)-f(x_i)|\frac{g_i^\lambda(x)}{\sum_{j=1}^N g_j^\lambda(x)}\le\sum_{i\in\{i:\|x-x_i\|_2<2h_X\}}\\ &|f(x)-f(x_i)|\frac{g_i^\lambda(x)}{\sum_{j=1}^N g_j^\lambda(x)}+\sum_{i\in\{i:\|x-x_i\|_2\ge 2h_X\}}\\ &|f(x)-f(x_i)|\frac{g_i^\lambda(x)}{\sum_{j=1}^N g_j^\lambda(x)}\le\omega(f,2h_X)+\sum_{i\in\{i:\|x-x_i\|_2\ge 2h_X\}}\\ &|f(x)-f(x_i)|\frac{g_i^\lambda(x)}{\sum_{j=1}^N g_j^\lambda(x)}=:\omega(f,2h_X)+\Delta_1.\end{aligned}$$

To estimate Δ_1, we rewrite Δ_1 as

$$\Delta_1=\left(\sum_{i\in\{i:2h_X\le\|x-x_i\|_2<3h_X\}}+\sum_{i\in\{i:3h_X\le\|x-x_i\|_2<4h_X\}}+\cdots+\sum_{i\in\{i:kh_X\le\|x-x_i\|_2<(k+1)h_X\}}+\cdots\right)\times|f(x)-f(x_i)|\frac{g_i^\lambda(x)}{\sum_{j=1}^N g_j^\lambda(x)}.$$

From the definition of h_X it follows that for given x there exists