Proceedings of the Institute of Mathematics and Mechanics,
National Academy of Sciences of Azerbaijan
Volume 47, Number 2, 2021, Pages 226–231
https://doi.org/10.30546/2409-4994.47.2.226

ON THE ERROR OF APPROXIMATION BY RBF NEURAL NETWORKS WITH TWO HIDDEN NODES

ARZU M-B. BABAYEV AND IBRAHIM K. MAHAROV

Abstract. We consider the problem of approximation of a continuous multivariate function by RBF neural networks with two hidden nodes in the uniform norm. We obtain a sharp lower bound estimate for the approximation error in terms of functionals generated by closed paths.

1. Introduction

Broomhead and Lowe [7] were the first to introduce the RBF (Radial Basis Function) neural networks. Such neural networks are known as universal approximators due to their remarkable performance in the problem of function approximation. Originally, the RBF networks were designed for data interpolation in a higher dimensional space. However, their applications are in wide area of engineering and they have been used as an important tool for function approximation, prediction, estimation, and system control. The main advantage of RBF neural networks is the simplicity of computation of network parameters. These networks are able to perform complex nonlinear mappings and provide a fast and robust learning mechanism without significant computational cost.

The set of RBF neural networks consists of the following functions (see, e.g., [3])

$$\sum_{i=1}^{m} w_i g\left(\frac{\|\mathbf{x}-\mathbf{c}_i\|}{\sigma_i}\right). \tag{1.1}$$

Here $m \in \mathbb{N}$ is the number of nodes in the hidden layer, $(w_1,...,w_m) \in \mathbb{R}^m$ is the vector of weights, $\mathbf{x} \in \mathbb{R}^d$ is an input vector, $\mathbf{c}_i \in \mathbb{R}^d$ and $\sigma_i \in \mathbb{R}$ are the centroid and smoothing factor (or width) of the i-th node, $1 \leq i \leq m$, respectively, $\|\mathbf{x}-\mathbf{c}_i\|$ is the Euclidean distance between $\mathbf{x}$ and $\mathbf{c}_i$, and $g : \mathbb{R} \to \mathbb{R}$ is the so-called activation function.

Note that the functions of the form $r(\|\mathbf{x}-\mathbf{c}\|)$ involved in the right hand side of (1.1) are called *radial basis functions.* In other words, a radial basis function is a multivariate function constant on the spheres $\|\mathbf{x}-\mathbf{c}\| = \alpha,\ \alpha \in \mathbb{R}$. These functions and their linear combinations arise naturally in many fields, especially in RBF neural networks (see, e.g., [10, 17, 18, 21, 22, 23, 24, 25, 26]).

2010 *Mathematics Subject Classification.* 41A30, 41A63, 92B20.
Key words and phrases. radial basis function, RBF neural network, path, extremal element, approximation error.

Let $f(\mathbf{x})$ be a given continuous function on some compact subset Q of $\mathbb{R}^d$ and $g(x)$ be any continuous function on $\mathbb{R}$. Consider the approximation of f from the following set of RBF neural networks

$$\mathcal{G} = \mathcal{G}(g, \mathbf{c}_1, \mathbf{c}_2) = \left\{ w_1 g\left(\frac{\|\mathbf{x} - \mathbf{c}_1\|}{\sigma_1}\right) + w_2 g\left(\frac{\|\mathbf{x} - \mathbf{c}_2\|}{\sigma_2}\right) : \ w_i, \sigma_i \in \mathbb{R}, \ i = 1, 2 \right\}. \tag{1.2}$$

Note that in (1.2) $\mathbf{c}_1$ and $\mathbf{c}_2$ are fixed, whileas the numbers $w_1, w_2, \sigma_1, \sigma_2$ variate. The approximation error is defined as follows

$$E(f) = E(f, \mathcal{G}) \stackrel{def}{=} \inf_{u \in \mathcal{G}} \|f - u\|,$$

where

$$\|f - u\| = \max_{\mathbf{x} \in Q} |f(\mathbf{x}) - u(\mathbf{x})|.$$

In this paper, we are interested in lower bound estimates for the approximation error $E(f)$. We show in the next section that $E(f)$ can be estimated below by using values of specifically constructed functionals at f.

2. The approximation error estimation

Suppose Q is a compact set in $\mathbb{R}^d$ and $\mathbf{c}_1, \mathbf{c}_2 \in \mathbb{R}^d$ are fixed points.

Definition 2.1. (see [5]) *A finite or infinite ordered set $p = (\mathbf{p}_1, \mathbf{p}_2, ...) \subset Q$ with $\mathbf{p}_i \neq \mathbf{p}_{i+1}$, and either $\|\mathbf{p}_1 - \mathbf{c}_1\| = \|\mathbf{p}_2 - \mathbf{c}_1\|, \|\mathbf{p}_2 - \mathbf{c}_2\| = \|\mathbf{p}_3 - \mathbf{c}_2\|$, $\|\mathbf{p}_3 - \mathbf{c}_1\| = \|\mathbf{p}_4 - \mathbf{c}_1\|, ...$ or $\|\mathbf{p}_1 - \mathbf{c}_2\| = \|\mathbf{p}_2 - \mathbf{c}_2\|, \|\mathbf{p}_2 - \mathbf{c}_1\| = \|\mathbf{p}_3 - \mathbf{c}_1\|$, $\|\mathbf{p}_3 - \mathbf{c}_2\| = \|\mathbf{p}_4 - \mathbf{c}_2\|, ...$ is called a path with respect to the centers $\mathbf{c}_1$ and $\mathbf{c}_2$.*

In the above definition, we alternate distances from two fixed points. Paths have many different variations. For example, instead of points, one can take two hyperplanes $\mathbf{a}^i \cdot \mathbf{x} = \alpha_i$, $i = 1, 2$, where " $\cdot$ " denotes the standard scalar product in $\mathbb{R}^d$, and alternate distances from these two hyperplanes. Certainly, in $\mathbb{R}^2$, hyperplanes turn into straight lines, thus one can talk about distances from straight lines. Paths with respect to two straight lines in $\mathbb{R}^2$ were first considered by Braess and Pinkus [6]. They showed that paths give geometric means of deciding if a set of points $\{\mathbf{x}^i\}_{i=1}^m \subset \mathbb{R}^2$ has the non-interpolation property for so called "ridge functions". For these functions and their various properties see [1, 6, 12, 16]. Ismailov and Pinkus [12] used paths with respect to two hyperplanes in $\mathbb{R}^d$ to solve the problem of interpolation on straight lines by ridge functions. Paths with respect to two hyperplanes $\mathbb{R}^d$ were also used in [4, 14, 15, 16]. If two straight lines in $\mathbb{R}^2$ are taken as the coordinate lines, then the corresponding set of points $(\mathbf{p}_1, \mathbf{p}_2, ...)$ turn into "bolts of lightning" (see, e.g., [2, 11, 20]). It is well known that the idea of bolts was first introduced by Diliberto and Straus [8] and played an essential role in problems of approximation by sums of univariate functions (see, e.g., [8, 9, 11, 19, 20]). Note that the name "bolt of lightning" is due to Arnold [2]. Ismailov [13] generalized paths to those with respect to a finite set of functions. Paths with respect to n arbitrarily fixed functions turned out to be very useful in problems of representation by linear superpositions.

In the sequel, we use the term "path" instead of the long expression "path with respect to the centers $\mathbf{c}_1$ and $\mathbf{c}_2$". A finite path $(\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_{2n})$ is said to be closed if $(\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_{2n}, \mathbf{p}_1)$ is also a path.

We associate a closed path $p = (\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_{2n})$ with the functional

$$G_p(f) = \frac{1}{2n} \sum_{k=1}^{2n} (-1)^{k+1} f(\mathbf{p}_k).$$

The following theorem is valid.

Theorem 2.1. *Let a compact set Q have closed paths. Then the approximation error $E(f)$ of a function $f \in C(Q)$ from the set of RBF neural networks $\mathcal{G}(g, \mathbf{c}_1, \mathbf{c}_2)$ obeys the following sharp lower bound estimate*

$$E(f) \geq \sup_{p \subset Q} |G_p(f)|,$$

where the sup *is taken over all closed paths.*

To prove this theorem we need the following lemma.

Lemma 2.1. *Let a compact set Q have closed paths. Then*

$$\sup_{p \subset Q} |G_p(f)| \leq \inf_{r \in \mathcal{D}} \|f - r\|, \tag{2.1}$$

where the sup is taken over all closed paths and

$$\mathcal{D} = \{r_1(\|\mathbf{x} - \mathbf{c}_1\| + r_2(\|\mathbf{x} - \mathbf{c}_2\| : \ r_i \in C(\mathbb{R}), \ i = 1, 2\}.$$

Moreover, inequality (2.1) is sharp, i.e. there exist functions for which (2.1) turns into equality.

Proof. First note that the functional $G_p(f)$ has the following obvious properties:

(a) If $r \in \mathcal{D}$, then $G_p(r) = 0$.

(b) $\|G_p\| \leq 1$ and if $p_i \neq p_j$ for all $i \neq j$, $1 \leq i, j \leq 2n$, then $\|G_p\| = 1$.

Let now p be a closed path of Q and r be any function from $\mathcal{D}$. Then by the linearity of G_p and properties (a) and (b),

$$|G_p(f)| = |G_p(f - r)| \leq \|f - r\|. \tag{2.2}$$

Since the left-hand and the right-hand sides of (2.2) do not depend on r and p respectively, it follows from (2.2) that

$$\sup_{p \subset Q} |G_p(f)| \leq \inf_{r \in \mathcal{D}} \|f - r\|. \tag{2.3}$$

Now we prove the sharpness of (2.1). By assumption Q has closed paths. Then Q has closed paths $p' = (\mathbf{p}'_1, ..., \mathbf{p}'_{2m})$ such that all points$\mathbf{p}'_1, ..., \mathbf{p}'_{2m}$ are distinct. On the other hand there exist continuous functions $g = g(\mathbf{x})$ on Q such that $g(\mathbf{p}'_i) = 1$, $i = 1, 3, ..., 2m - 1$, $g(\mathbf{p}'_i) = -1$, $i = 2, 4, ..., 2m$ and $-1 < g(\mathbf{x}) < 1$ elsewhere. For such functions we have

$$G_{p'}(g) = \|g\| = 1 \tag{2.4}$$

and

$$\inf_{r\in\mathcal{D}} \|g-r\| \leq \|g\|, \tag{2.5}$$

where the last inequality follows from the fact that $0 \in \mathcal{D}$. From (2.3)-(2.5) it follows that

$$\sup_{p\subset Q} |G_p(g)| = \inf_{r\in\mathcal{D}} \|g-r\|.$$

We have proved the sharpness of (2.1) and hence the lemma.

Proof of Theorem 2.1. Note that in the definition of $\mathcal{G}(g, c_1, c_2)$ each term $w_i g\left(\frac{\|\mathbf{x}-\mathbf{c}_i\|}{\sigma_i}\right)$, $i = 1, 2$, can be considered as a function $g_i\left(\|\mathbf{x}-\mathbf{c}_i\|\right)$. Here g_i depends on the parameters w_i and σ_i. Thus we see that an element $v \in \mathcal{G}(g, c_1, c_2)$ also belongs to the class $\mathcal{D}$; therefore, $\mathcal{G}(g, c_1, c_2) \subset \mathcal{D}$. This means that

$$E(f) = \inf_{u\in\mathcal{G}} \|f-u\| \geq \inf_{r\in\mathcal{D}} \|f-r\|. \tag{2.6}$$

It follows from (2.6) and Lemma 2.1 that

$$E(f) \geq \sup_{p\subset Q} |G_p(f)|, \tag{2.7}$$

where the sup is taken over all closed paths in Q.

The sharpness of (2.7) can be proved similarly as the sharpness of the inequality (2.1). Indeed, take a closed path $p = (\mathbf{p}_1, ..., \mathbf{p}_{2k}) \subset Q$ such that all points $\mathbf{p}_1, ..., \mathbf{p}_{2k}$ are distinct. Choose a continuous function $f_0 = f_0(\mathbf{x})$ on Q such that $f_0(\mathbf{p}_i) = 1$, $i = 1, 3, ..., 2k-1$, $f_0(\mathbf{p}_i) = -1$, $i = 2, 4, ..., 2k$ and $-1 < f(\mathbf{x}) < 1$ elsewhere. For such functions we have

$$G_p(f_0) = \|f_0\| = 1. \tag{2.8}$$

Now it follows from (2.7), (2.8) and the fact $\|f_0\| \geq E(f_0)$ that

$$G_p(f_0) \geq E(f_0) \geq \sup_{p\subset Q} |G_p(f_0)|.$$

Therefore,

$$E(f_0) = \sup_{p\subset Q} |G_p(f_0)|.$$

The theorem has been proven.

The following theorem is based on Theorem 2.1.

Theorem 2.2. *Let Q be a compact set in $\mathbb{R}^d$ and $\mathbf{c}_1, \mathbf{c}_2 \in \mathbb{R}^d$. Let $\mathbf{x}_1, \mathbf{x}_2 \in S_1 \cap S_2 \cap Q$, where S_1 and S_2 are spheres centered at $\mathbf{c}_1$ and $\mathbf{c}_2$, respectively. Then the approximation error $E(f)$ of a function $f \in C(Q)$ from the set of RBF neural networks $\mathcal{G}(g, \mathbf{c}_1, \mathbf{c}_2)$ obeys the following two-sided estimates*

$$\frac{1}{2}\left|f(\mathbf{x}_1) - f(\mathbf{x}_2)\right| \leq E(f) \leq \|f\|.$$

Proof. If $\mathbf{x}_1, \mathbf{x}_2 \in S_1 \cap S_2 \cap Q$, then $\|\mathbf{x}_1 - \mathbf{c}_1\| = \|\mathbf{x}_2 - \mathbf{c}_1\|$ and $\|\mathbf{x}_1 - \mathbf{c}_2\| = \|\mathbf{x}_2 - \mathbf{c}_2\|$. Thus $p_0 = (\mathbf{x}_1, \mathbf{x}_2)$ is a closed path (with respect to the centers $\mathbf{c}_1$ and $\mathbf{c}_2$) in Q. By Theorem 2.1,

$$E(f) \geq \sup_{p\subset Q} |G_p(f)| \geq |G_{p_0}(f)| = \frac{1}{2}|f(\mathbf{x}_1) - f(\mathbf{x}_2)|. \quad (2.9)$$

On the other hand, since the zero function belongs to $\mathcal{G} = \mathcal{G}(g, \mathbf{c}_1, \mathbf{c}_2)$, we can write that

$$E(f) = \inf_{u\in\mathcal{G}} \|f - u\| \leq \|f - 0\| = \|f\|. \quad (2.10)$$

Combining (2.9) and (2.10) we finally obtain that

$$\frac{1}{2}|f(\mathbf{x}_1) - f(\mathbf{x}_2)| \leq E(f) \leq \|f\|.$$

The theorem has been proved.

References

[1] R.A. Aliev, A.A. Asgarova, V.E. Ismailov, On the Hölder continuity in ridge function representation, *Proc. Inst. Math. Mech. Natl. Acad. Sci. Azerb.* **45** (2019), no. 1, 31-40.

[2] V.I. Arnold, On functions of three variables, *Dokl. Akad. Nauk SSSR* **114** (1957), 679-681; English transl., *Amer. Math. Soc. Transl.* **28** (1963), 51-54.

[3] C. Arteaga, I. Marrero, Universal approximation by radial basis function networks of Delsarte translates, *Neural Networks* **46** (2013), 299-305.

[4] A.Kh. Asgarova, A.M-B. Babayev, I.K. Maharov, On the error of approximation by ridge functions with two fixed directions, *Tbilisi Math. J.* **10** (2017), no. 2, 111-120.

[5] A.Kh. Asgarova, A.M-B. Babayev, I.K. Maharov, On the error of approximation by radial basis functions with fixed centers, *Trans. Natl. Acad. Sci. Azerb. Ser. Phys.-Tech. Math. Sci.* **38** (2018), no. 1, 22-29.

[6] D. Braess, A. Pinkus, Interpolation by ridge functions, *J. Approx. Theory* **73** (1993), 218-236.

[7] D.S. Broomhead, D. Lowe, Multivariable function interpolation and adaptive networks, *Complex Systems* **2** (1988), 321-355.

[8] S.P. Diliberto, E.G. Straus, On the approximation of a function of several variables by the sum of functions of fewer variables, *Pacific J. Math.* **1** (1951), 195-210.

[9] M.v. Golitschek, W.A. Light, Approximation by solutions of the planar wave equation, *Siam J. Numer. Anal.* **29** (1992), 816-830.

[10] Q.P. Ha, H.Wahid, H. Duc, M. Azzi, Enhanced radial basis function neural networks for ozone level estimation, *Neurocomputing* **155** (2015), 62-70.

[11] S.Ja. Havinson, A Chebyshev theorem for the approximation of a function of two variables by sums of the type $\varphi(x) + \psi(y)$, *Izv. Acad. Nauk. SSSR Ser. Mat.* **33** (1969), 650-666; English tarnsl. *Math. USSR Izv.* **3** (1969), 617-632.

[12] V.E. Ismailov, A. Pinkus, Interpolation on lines by ridge functions, *J. Approx. Theory* **175** (2013), 91-113.

[13] V.E. Ismailov, A note on the representation of continuous functions by linear superpositions, *Expo. Math.* **30** (2012), 96-101.

[14] V.E. Ismailov, On the proximinality of ridge functions, *Sarajevo J. Math.* **5**(17) (2009), no. 1, 109-118.

[15] V.E. Ismailov, A note on the equioscillation theorem for best ridge function approximation, *Expo. Math.* **35** (2017), no. 3, 343-349.

[16] V.E. Ismailov, Approximation by sums of ridge functions with fixed directions, (Russian) translated from *Algebra i Analiz* **28** (2016), no. 6, 20-69; English transl. *St. Petersburg Math. J.* **28** (2017), no. 6, 741-772.
[17] C.-N. Ko, Reinforcement radial basis function neural networks with an adaptive annealing learning algorithm, *Appl. Math. Comput.* **221** (2013), 503-513.
[18] S. Lin, Linear and nonlinear approximation of spherical radial basis functions networks, *J. Complexity* **35** (2016), 86-101.
[19] D.E. Marshall, A.G.O'Farrell, Uniform approximation by real functions, *Fund. Math.* **104** (1979), 203-211.
[20] D.E. Marshall and A.G.O'Farrell, Approximation by a sum of two algebras. The lightning bolt principle, *J. Funct. Anal.* **52** (1983), 353-368.
[21] G.A. Montazer, D. Giveki, An improved radial basis function neural network for object image retrieval, *Neurocomputing* **168** (2015), 221-233.
[22] H. Muzhou, H. Xuli, The multidimensional function approximation based on constructive wavelet RBF neural network, *Appl. Soft Comput.* **11** (2011), 2173-2177.
[23] Y. Xie, Y. Xing, A prediction method for the wax deposition rate based on a radial basis function neural network, *Petroleum* **3** (2017), 237-241.
[24] S. Yang, T.O. Ting, K.L. Man, S.-U. Guan, Investigation of neural networks for function approximation, *Procedia Computer Science* **17** (2013), 586-594.
[25] S.-H. Yoo, S.-K. Oh, W. Pedrycz, Optimized face recognition algorithm using radial basis function neural networks and its practical applications, *Neural Networks* **69** (2015), 111-125.
[26] H. Yuan, X. Shen, X. Chen, Prediction of fishing ground based on RBF neural network, *Procedia Engineering* **15** (2011), 3240-3244.

Arzu M-B. Babayev
Institute of Mathematics and Mechanics, NAS of Azerbaijan, Baku, AZ 1141, Azerbaijan
E-mail address: `arzumb.babayev@gmail.com`

Ibrahim K. Maharov
Institute of Mathematics and Mechanics, NAS of Azerbaijan, Baku, AZ 1141, Azerbaijan
E-mail address: `ibrahimmaharov@gmail.com`

Received: November 28, 2020; Accepted: April 15, 2021