Generalized-Multiquadric Radial Basis Function Neural Networks (RBFNs) with Variable Shape Parameters for Function Recovery

Sayan Kaennakham[a,1], Pichapop Paewpolsong[a], Natdanai Sriapai[a] and Sunisa Tavaen[a]
[a] *School of Mathematics, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima, Thailand*

Abstract. After being introduced to approximate two-dimensional geographical surfaces in 1971, the multivariate radial basis functions (RBFs) have been receiving a great amount of attention from scientists and engineers. In 1987 the idea was extended into the construction of neural networks corresponding to the beginning of the era of artificial intelligence, forming what is now called 'Radial Basis Function Neural Networks (RBFNs)'. Ever since, RBFNs have been developed and applied to a wide variety of problems; approximation, interpolation, classification, prediction, in nowadays science, engineering, and medicine. This also includes numerically solving partial differential equations (PDEs), another essential branch of RBFNs under the name of the 'Meshfree/Meshless' method. Amongst many, the so-called 'Multiquadric (MQ)' is known as one of the mostly-used forms of RBFs and yet only a couple of its versions have been extensively studied. This study aims to extend the idea toward more general forms of MQ. At the same time, the key factor playing a very crucial role for MQ called 'shape parameter' (where selecting a reliable one remains an open problem until now) is also under investigation. The scheme was applied to tackle the problem of function recovery as well as an approximation of its derivatives using six forms of MQ with two choices of the variable shape parameter. The numerical results obtained in this study shall provide useful information on selecting both a suitable form of MQ and a reliable choice of MQ-shape for further applications in general.

Keywords. Generalized-Multiquadric, Radial Basis Function Neural Networks (RBFNs), Variable Shape Parameters, Function Recovery

1. Introduction

Radial Basis Functions (RBFs), φ, are commonly found as multivariate functions whose values are dependent only on the distance from the origin. This means that $\varphi(\mathbf{x}) = \varphi(r) \in \mathbb{R}$ with $\mathbf{x} \in \mathbb{R}^n$ and $r \in \mathbb{R}$, or, in other words, on the distance from a point of a given set $\{\mathbf{x}_j\}$, and $\varphi(\mathbf{x} - \mathbf{x}_j) = \varphi(r_j) \in \mathbb{R}$. Here, r_j is the Euclidean distance. As illustrated in Figure 1, RBF networks (RBFNs) broadly consist of three layers; 1) *Input*

---

[1] Corresponding Author, Sayan Kaennakham, School of Mathematics, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima, Thailand; E-mail: sayan_kk@g.sut.ac.th.

layer – This layer is for independent variables to be imported with their corresponding values. 2) Hidden layer – This is where all the imported data is subjected to the constraints of each neuron. 3) Output later – This layer summaries in a linear combination manner the semi-output variables with their associated weights (w1, w2, ..., wm) obtained from the previous layer.

The very first problem of interpolation in multidimensional space that was approached by the concept of RBFNs is that documented by Powel [1]. This kind of problem requires as many centres as data points making it challenging for convectional interpolation schemes. With the capability of generating any non-linear mappings between stimulus and response, RBFNs have been developed and applied in a wide range of applications from machine learning to solving partial differential equations (PDE) [2]. Until nowadays, it can be summarized in general that researches on RBFNs are grouped into three branches [3]: i) kernels formation and development, ii) learning mechanism or weight optimization and judgments, and iii) the areas of application, of which the detail shall be provided below.
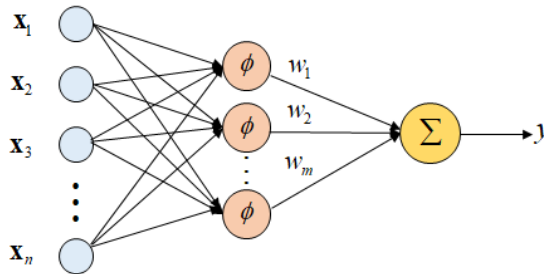

Figure 1. Typical structure of RBFNs.

Some successful applications are the approximation of function in 1-D and 2-D using polynomial with RBFNs [4], the solution to regulator equations based on a class of RBFNs [5],the use of DE- RBFNs for classifying weblog dataset [6], the use of RBFNs optimization for support vector machine classifiers [7], the application of Hermite collocation scheme for coupled PDEs [8], and the application of the modified MQ-RBFNs [9].

In all aspects of the application of RBFNs mentioned so far, searching for a suitable and reliable choice of a shape parameter, embedded in most RBF forms, still remains a great challenge. In this work, we numerically investigate the effectiveness of two recently proposed choices in a variable manner of the parameter as well as focus on a wider range of versions of multiquadric basis function, called 'generalized MQ'. One and two-dimensional problems are tackled; both the functions themselves and their derivatives. This work is organised as follows. Section 2 provides a short mathematical derivation of the RBFNs before the detail of generalized MQ is provided in section 3 with the two versions of variable shape. Section 4 then demonstrates the main results of the experiment before all the crucial findings are listed in section 5.

2. Function Approximation by RBFNs

It starts with considering the interpolation of a multivariate function $f : \Omega \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^n$, from a set of sample values $\left\{ f(\mathbf{x}_j) \right\}_{j=1}^{N}$ on a discrete set $X = \left\{ \mathbf{x}_j \right\}_{j=1}^{N} \subset \Omega$ (later referred to as 'centres'). These multivariate functions can well be reproduced if appropriate linear combinations of univariate interpolation functions with the Euclidean norm $\| \cdot \|_2$ are used. This process can be carried out using translates $\Phi(\mathbf{x} - \mathbf{x}_j)$ of a continuous real-valued function Φ defined on $\mathbb{R}$, and by letting Φ be radially symmetric; i.e.,

$$\Phi(\mathbf{x}) := \varphi\left(\| \mathbf{x} \|_2 \right) \tag{1}$$

with a continuous function φ on $\mathbb{R}_0^+$. In general science and engineering literature, φ is often called a radial basis function with centres $\left\{ \mathbf{x}_j \right\}_{j=1}^{N}$ and Φ is the associated kernel.

Interpolants F to f can be constructed as.

$$F(\mathbf{x}) = \sum_{j=1}^{N} a_j \varphi\left(\| \mathbf{x} - \mathbf{x}_j \|_2 \right) \tag{2}$$

with real coefficients, $\left\{ a_j \right\}_{j=1}^{N}$ which can be determined by the following interpolating condition.

$$F(\mathbf{x}_i) = f(\mathbf{x}_i), \tag{3}$$

for all $i = 1, 2, ..., N.$ and it leads to the following form.

$$f(\mathbf{x}_i) = F(\mathbf{x}_i) = \sum_{j=1}^{N} a_j \varphi\left(\| \mathbf{x}_i - \mathbf{x}_j \|_2 \right), \; 1 \le i \le N, \tag{4}$$

Hence, what comes next is a system of linear equations with $\left\{ a_j \right\}_{j=1}^{N}$ being the unknowns, expressed as follows.

$$\varphi\boldsymbol{\alpha} = \mathbf{f} \tag{5}$$

where $\varphi = \left[\varphi_{ij} \right]_{N \times N}$, $\boldsymbol{\alpha} = \left[a_1, ..., a_N \right]^T$, and $\mathbf{f} = \left[f(\mathbf{x}_1), ..., f(\mathbf{x}_N) \right]^T$. The existence of the above system depends highly on the choice of RBF being used (will be more discussed later). Once the coefficient matrix is obtained, the solution calculation process

can then proceed as follows. For unknown locations $\hat{X}=\left\{\hat{\mathbf{x}}_j\right\}_{j=1}^{\hat{N}} \subset \Omega$ and $X \cap \hat{X}=\phi$, the approximate values of the corresponding function $\hat{f}\left(\hat{\mathbf{x}}_j\right)$ can be carried out using the expression below.

$$\hat{f}\left(\hat{\mathbf{x}}_j\right) \approx F(\hat{\mathbf{x}}_j) = \sum_{i=1}^{N} a_j \varphi\left(\left\|\hat{\mathbf{x}}_j - \mathbf{x}_i\right\|_2, \varepsilon\right) \tag{6}$$

for all $j = 1, 2, ..., \hat{N}$. Note that the shape parameter ε is added to the equation above to emphasize the objective of this work which deals with the effect of this crucial parameter when the popular multiquadric RBF is utilized.

This similar assumption of approximation using RBFNs can be extended to the approximation of the n-order of function's derivative, provided that the basis function is also n-order differentiable. It is done via the following relation.

$$\frac{\partial^{(n)}}{\partial x_k}\left[\hat{f}\left(\hat{\mathbf{x}}_j\right)\right] \approx \frac{\partial^{(n)}}{\partial x_k}\left[F(\hat{\mathbf{x}}_j)\right] = \sum_{i=1}^{N} a_j \frac{\partial^{(n)}}{\partial x_k} \varphi\left(\left\|\hat{\mathbf{x}}_j - \mathbf{x}_i\right\|_2, \varepsilon\right) \tag{7}$$

The next section provides all components used and addressed in this investigation before some demonstrating 1D and 2D numerical tests are carried out in section 4.

3. Generalized MQ and It's Shape Choices

The main focus of this work is on one of the popular form of basis function known as 'multiquadric (MQ)' in its generalized form defined as follows.

$$\varphi_j(x,y) = \left(\varepsilon^2 + r_j^2\right)^\beta = \left(\varepsilon^2 + (x - x_j)^2 + (y - y_j)^2\right)^\beta \tag{8}$$

where $\beta = ..., -3/2, -1/2, 1/2, 3/2, ...$. These values give direct effect on the singularity of the interpolation matrix and, unavoidably, the effectiveness of the methods' performance. Different values yield different shape curves depicted in Figure 2 (a).
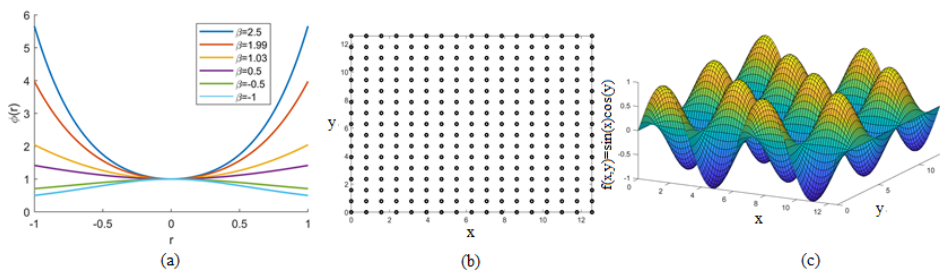

Figure 2. (a) Generalized MQ with six values of $\beta's$, (b) distribution of centres, and (c) two-dimensional test function.

One crucial factor affecting the method's performance is the shape parameter. Choices available in literature can be categorized into three forms; constant/fixed, variable, and iterative-based. This work focuses on two nonlinear-variable shape strategies (illustrated in Figures 3(a)-3(b)).

- Strategy-1 (*Stg-1*): by Nojavan et.al. (2017) [10], defined as.

$$\varepsilon_j = (c_{min} + (c_{max} - c_{min})\exp(-j))^{-1} \qquad (9)$$

- Strategy-2 (*Stg-2*): by Xiang et.al. (2012) [11], defined as.

$$\varepsilon_j = c_{min} + (c_{max} - c_{min})\sin(j) \qquad (10)$$

with $j = 1, 2, ..., N$. Apart from these chosen forms of shape selection strategies, those nicely documented in [12] and [13] are also hightly recommented for interested readers.
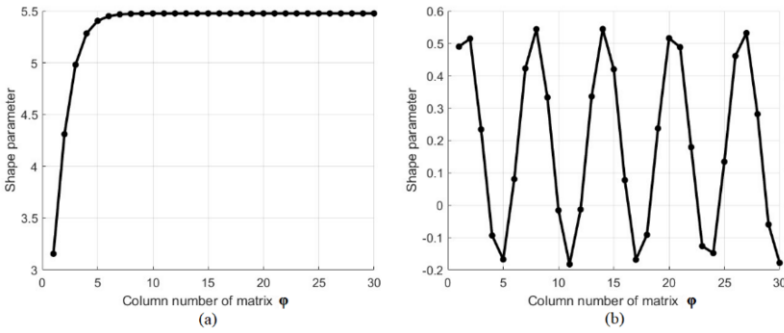

Figure 3. Multiquadrics' shape determining strategies; (a) *Stg-1*, and (b) *Stg-2* both using $c_{min} = 1/\sqrt{N}$, $c_{max} = 3/\sqrt{N}$, and $N = 30$.

4. Numerical Experiments and Results

Results obtained from each case are evaluated using the mean absolute error norms defined as follows.

$$Mean\ Absolute\ Error = \frac{1}{\hat{N}}\left(\sum_{i=1}^{\hat{N}}\left|u_i^{exact}(x, y) - u_i^{approx.}(x, y)\right|\right) \qquad (11)$$

To cover all subcases involved, a great number of simulations were performed on a machine with Windows 10 (64bit), Intel Corei7-11800H CPU (2.30GHz ,26MB L3 4.60

GHz), 32 GB DDR4 3200MHz. Two main cases are under investigation and presented here.

4.1. One-Dimensional Case

The first case deals with the approximation of a function in one dimension expressed below.

$$f(x) = e^{x^3} + \cos(2x) \tag{12}$$

To compare the overall accuracy, those shape values stated in the work of Maggie [14] were also tested. The total of 25 nodes uniformly distributed over a $-1 \le x \le 1$ domain was considered where the function's values at 17 equally-spaced nodes were numerically approximated. The results were obtained from all six forms of MQ with two variable shapes (**Stg-**1 and **Stg-**2) together with fixed values are presented in Table 1. It can be shown in the table that all cases performed reasonably well with not much of a significant difference. One interesting finding worth mentioning is that those values obtained by using a fixed value were found the best ones. This might be attributed to the fact that Maggie [14] had actually performed a tremendous amount of experiments to find the best value of shape before utilizing that value for the actual experiment.

Table 1. Mean absolute errors measured for approximation of the function and its derivative for all forms of generalized MQ when using different variable shape forms (compared with a fixed form from literature).

Type	Shape	$\beta = 2.50$	$\beta = 1.99$	$\beta = 1.03$	$\beta = 0.50$	$\beta = -0.50$	$\beta = -1.00$
	Stg-1	28.70E-05	12.00E-04	71.60E-05	27.00E-04	11.66E-05	40.57E-05
$f(x)$	*Stg-2*	61.07E-07	15.18E-06	42.77E-06	48.52E-06	42.48E-05	16.00E-04
	Maggie [14]	16.00E-04	16.00E-04	59.44E-05	89.55E-06	74.68E-07	54.15E-07
	Stg-1	22.00E-04	70.00E-04	35.00E-04	81.00E-04	12.00E-04	30.00E-04
$f'(x)$	*Stg-2*	29.36E-05	71.61E-05	20.00E-04	23.00E-04	17.80E-03	66.00E-03
	Maggie [14]	33.00E-04	86.00E-04	20.00E-04	21.20E-05	10.05E-05	91.73E-06

4.2. Two-Dimensional Case

A more complex problem is now visited and a two-dimensional function is addressed, defined as follows.

$$f(x, y) = \cos(y) \cdot \sin(x) \tag{13}$$

This is defined on a $[0, 4\pi] \times [0, 4\pi]$ domain where the corresponding centres distribution is depicted in Figure 2(b), and its surface is being visualized in Figure 2(c). To focus on the two main variable shape strategies, the ones from the literature are omitted here. Five densities of centres were tested for each strategy to observe the overall

behavior of the variable when dealing with larger sizes of datasets. The set of 17×17 equally-spaced is used for the target locations for approximation. As depicted in Figures 4-5-6, all three forms; the function itself, its x-direction derivative, and its y-direction derivative, reveal the same trends of accuracy. Those functions' approximation produced by shape strategy 1 (or **Stg-1**) is found to be highly sensitive to the increase in centres (N) while **Stg-2** is not. This could be attributed to the fact that **Stg-1** is in an exponential curve where it can easily be affected by the pre-judged values of C_{min} and C_{max}, whose values depend on N. The best accuracy is found for **Stg-1** when $\beta = -0.50$ and $N = 225$ are used where the comparatively worst results are revealed when $\beta = 2.50$ and $N = 625$ are used, see Figures 4(a), 5(a), and 6(a). The opposite scenario is found in the use of **Stg-2** where $\beta = 2.50$ is clearly seen to outperform the rest for all three targeted function forms, see Figures 4(b), 5(b), and 6(b). As can be anticipated, **Stg-2** produces better solutions, for all values of $\beta's$, when more supporting centres are involved, indicating the benefits one can achieve when deploying the strategy for larger datasets. It can also be seen from these same figures that the fastest reduction in error is obtained from $\beta = 2.50$ whereas the slowest one is from $\beta = -1.00$.
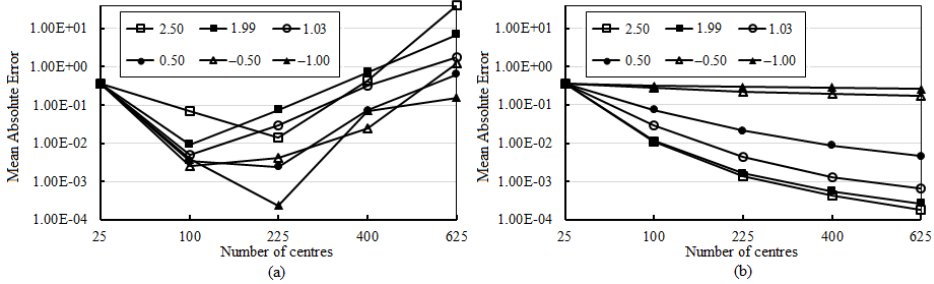

Figure 4. Mean absolute errors measured at different numbers of centres for approximation of $f(x, y)$; (a) using **Stg-1**, and (b) using **Stg-2**.
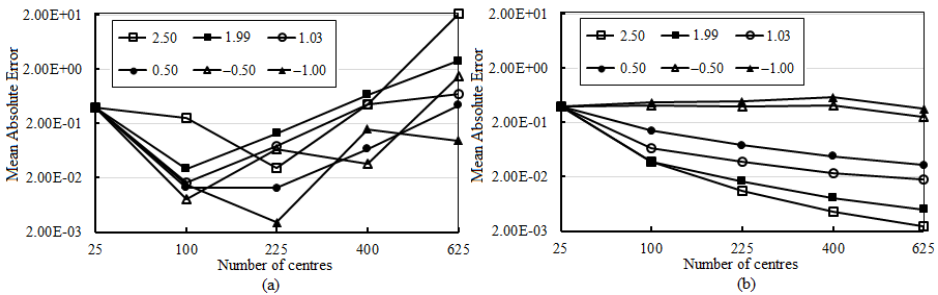

Figure 5. Mean absolute errors measured at different numbers of centres for approximation of $\partial f / \partial x$; (a) using **Stg-1**, and (b) using **Stg-2**.
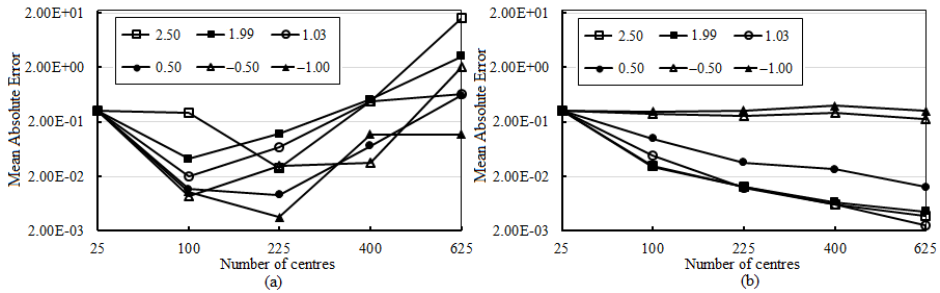
Figure 6. Mean absolute errors measured at different numbers of centres for approximation of $\partial f / \partial y$; (a) using ***Stg-1***, and (b) using ***Stg-2***.

5. Conclusions

Emerging as an alternative numerical tool for approximating and recoverying functions and their derivatives, multiquadric radial basis function neural networks (MQ-RBFNs) are under investigation in this work. Six generalized forms of MQ were numerically applied using two popular choices of shape parameters. It was found that the exponential variable shape (Stg-1) is highly sensitive to the number of centres for all forms of MQ used. On the other hand, the trigonometric variable shape (Stg-2) provides a monotone of reduction in errors with the increase of centres indicating a more reliable aspect for larger datasets. With the effectiveness in approximating derivatives of a function being also under the study, Stg-2 is seen promising with a great level of accuracy when the form $\left(\varepsilon^2 + r^2 \right)^{5/2}$ of MQ is used. This could well have a great effect in the era of solving partial differential equations which shall remain one of our future investigations.

References

[1]   Powell MJD. Radial basis functions for multivariable interpolation: A Review. In: Mason JC, Cox MG, editors. Algorithms for the Approximation. Oxford: Clarendon Press; 1987. p. 143-67.

[2]   Park J, Sandberg IW. Universal approximation using radial basis function networks. Neural Comput. 1991 June;3(2):246-57.

[3]   Beatson RK, Cherrie JB, Mouat CT. Fast fitting of radial basis functions: Methods based on preconditioned GMRES iteration. Advances in Math Comp. 1999 Nov;11(2):253-70. Available from: DOI:10.1023/A:1018932227617.

[4]   Bellil W, Amar CB, Alimi AM. Comparison between beta wavelets neural networks, RBF neural networks and polynomial approximation for 1D, 2D functions approximation. Trans. Eng. Comp. Tech. 2008. Available from: DOI:10.5281/zenodo.1056514.

[5]   Zhou G, Wang C, Su W. Nonlinear output regulation based on RBF neural network approximation. Proceeding of the *2005 International Conference on Control and Automation*; 2005 June 26-29; Budapest, Hungary. IEEE; 2005. p. 679-84. Available from: DOI:10.1109/ICCA.2005.1528210.

[6]   Dash CSK, Behera AK, Pandia MK, Dehuri S. Neural networks training based on differential evolution in radial basis function networks for classification of web logs. Proceeding of the International Conference on Distributed Computing and Internet Technology; 2013; Berlin, Heidelberg. Springer; 2013. p. 183-94. Available from: DOI:10.1007/978-3-642-36071-8_14.

[7]   Thurnhofer-Hemsi K, López-Rubio E, Molina-Cabello MA, Najarian K. Radial basis function kernel optimization for support vector machine classifiers. arXiv:2007.08233[Preprint]. 2020[cited 2020 Jul 16]. Available from: https://arxiv.org/abs/2007.08233.

[8]   Chuathong N, Kaennakham S. Numerical solution to coupled Burgers' equations by gaussian-based hermite collocation scheme. J. Appl. Math. 2018 Sep;2018:1-18. Available from: DOI:10.1155/2018/3416860.

[9]   Kaennakham S, Chanthawara K. Numerical solution to nonlinear transient coupled-PDE by the modified multiquadric meshfree method. Rom. J. Phys. 2021 Feb;66(108):1-14.

[10]   Nojavan H, Abbasbandy S. Allahviranloo T. Variable shape parameter strategy in local radial basis functions collocation method for solving the 2D nonlinear coupled Burgers' equations. Mathematics. 2017 Jul;5(3):1-21. Available from: DOI:10.3390/math5030038.

[11]   Xiang S, Wang K, Ai Y, Sha Y, Shi H. Trigonometric variable shape parameter and exponent strategy for generalized multiquadric radial basis function approximation. Appl. Math. Model. 2012 May;36(5):1931-8.

[12]   Cheng AHD, Multiquadric and its shape parameter-A numerical investigation of error estimate, condition number, and round-off error by arbitrary precision computation, Eng. Anal. with Boundary Elements. 2012: 36(2): 220-239. Available from: DOI: 10.1016/j.enganabound.2011.07.008.

[13]   Chen W, Hong Y, Lin J, The sample solution approach for determination of the optimal shape parameter in the Multiquadric function of the Kansa method, Computers and Mathematics with Applications. 2018: 75(8): 2942-2954. Available from: DOI:10.1016/j.camwa.2018.01.023.

[14]   Chenoweth ME, Sarra SA. A numerical study of generalized multiquadric radial basis function interpolation. SIURO. 2009;2(2):58-70. Available from: DOI:10.1137/09S01040X.