Function Estimation by Radial basis function network

Setareh Roshan
Computer Department
Shahid Rajaee Teacher Training University
Tehran, Iran
setare.roshan1996@gmail.com

***Abstract*— In these days, artificial networks use in classification and regression. There are many kinds of networks, but here I introduced Radial basis function networks, or RBFs. RBFs are a commonly used type of artificial neural network for function approximation problems. In this study I will change the number of centers and variance and conclude that there is an optimal variance in any problem. Another important result is more centers is better, but number of centers and variance have reverse relationship.**

Keywords—regression, prediction, radial basis function, neural network, RBF network

I. Introduction

Neural networks are a series of algorithms that mimic the operations of a human brain to recognize relationships between vast amounts of data. They are used in a variety of applications in financial services, from forecasting and marketing research to fraud detection and risk assessment.

In the field of mathematical modeling, a radial basis function network is an artificial neural network that uses radial basis functions as activation functions. The output of the network is a linear combination of radial basis functions of the inputs and neuron parameters. Radial basis function networks have many uses, including function approximation, time series prediction, classification, and system control. [1] [2] [3]

In this study we will focus on radial basis function (RBF). RBF is a popular kernel function used in various kernelized learning algorithms. Radial Basis Function Networks. Radial basis function (RBF) networks are a commonly used type of artificial neural network for function approximation problems. Depending on the case, it is typically observed that the RBF network required less time to reach the end of training compared to MLP.

RBF uses kernels instead of hidden neurons. In this study, I want to know how these centers, variance, and epochs effect on the performance of a regression target.

II. Methods

In this experiment, I used (1)

Radial Basis Function to estimate the outcome of Eq.(1). Topology of this network is 1:centers:1, the number of centers is variable, as it shows in Figure .

$$f(x) = 0.05x^3 - 0.2x^2 - 3x + 20$$

(1)

Radial Basis Function

The training of the RBF model is terminated once the calculated error reached the desired values or number of training iterations already was completed. An RBF network with a specific number of nodes in its hidden layer is chosen (**Figure**). A Gaussian function Eq.(2) is used as the transfer function in computational units [4].

$$hidden\ layer\ output_j = \exp\left(-\frac{||sample - mean_{cluster_i}||}{2\sigma^2_{cluster_i}}\right)$$

(2)

For each sample in training set we compute its hidden layer output then we will compute output like we did in MLP. In computing hidden layer output we don't have weights. In contrast in output layer, we have weights (random initial weights). For computing output, we use Eq.(3):

$$output_z = 1/(1 + \exp(-hidden\ layer\ output_j * weight_{z,j}))$$

(3)

For each $output_z$ we should multiple all hidden layer output that it's connected to $output_z$ with their weights and perform a sigmoid function on them. At last, we will compute an error (for each iteration) then with this error we will be able to update the weights (Eq. (4)).

$$new\ weight_{z,k} = old\ weight_{z,j} + \eta * error_z * (hidden\ layer\ output_j - output_z)$$

(4)

Which η is learning rate.

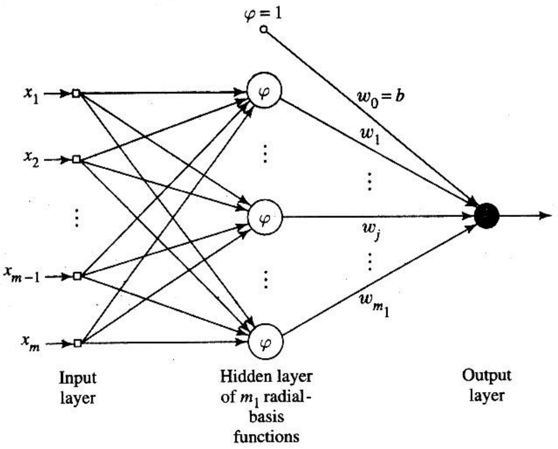

***Figure RBF neural network**. In this model there are three kernels (centers), m output neurons and an output neuron.*

Consider, in output layer there are weights, and in input layer there are no weights.

III. Experimental Results

In the first experiment I considered 21 numbers in -10 and 10 with step 1 (from now on I show it as: [-10,10,1]), epoch equal to 10000, and variance = 1. My network topology is 1:21:1, which means one input 21 centers and 1 output. According to Figure 1 the network fit the train data (which I considered in -10 to 10) properly.

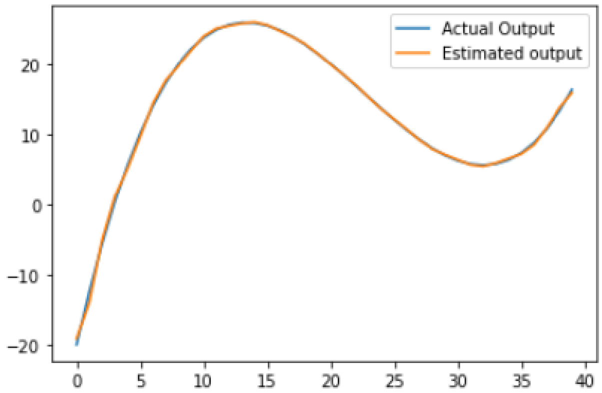

Figure 1 ***Fitted Trained Data****. Orange diagram represents fitted trained data, and blue diagram is actual targets. I fitted this data by 21 centers and 10000 epochs, also variance equal to 1.*

Interestingly, RBF cannot approximate values out of its range. I considered [-15,15,1] for test data which its result is in Figure 2 and it supports the idea that RBF is not a proper estimator for data which are out of its trained range.

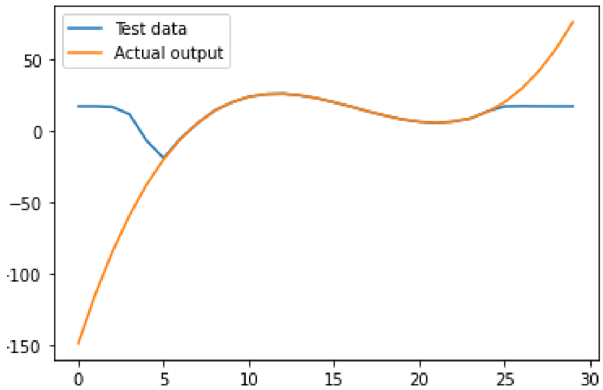

Figure 2 ***Fitted Test Data****. Orange diagram represents actual output values of test data, and blue diagram shows estimated output of test data by RBF.*

For comparing different amount of variance with 21 centers, I changed the variance 0.1, 1, and 200. According to Figure 1, Figure 2, and Figure 3 variance is important, high variance causes the model to fail proper approximation for both the test and train data. More important low variance, also cause this but not as bad as high variances, so in this model variance 1 had an acceptable estimation.

In another experiment, I fixed number of centers to 7 ([-10,10,5]) then I changed the variance (1, and 6). I can conclude form this experiment that in small center numbers it's for performance to choose wider widths (Figure 4Figure 5). Figure 4A model didn't approximate properly both test, and train data, but Figure 4B had a better performance. Then I fixed number of centers to 4 and changed the variance results showed in Figure 5 (1, and 6). The result was same as 7 centers.

IV. Conclusion

We can conclude that more centers have better accuracy, but amount of variance is challenging. If we choose variance small or large model performance would decrease, but when we choose an optimal variance model will accuracy is acceptable. In some cases, we don't have enough centers so in these cases it's better to choose variance wider.

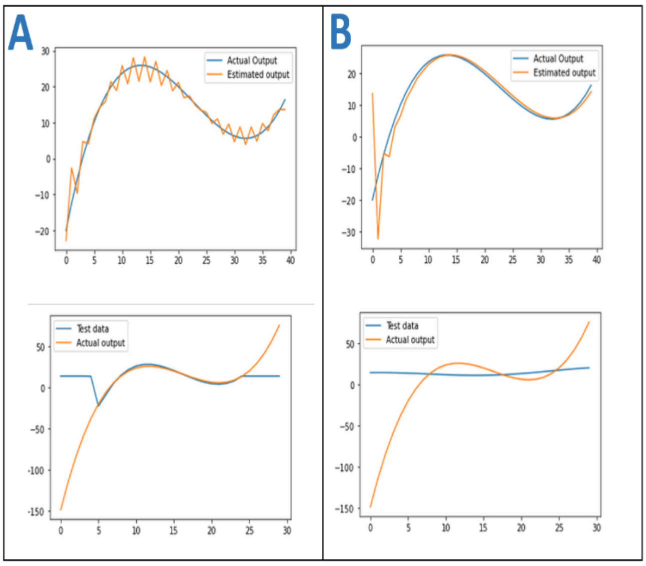

Figure 3 ***Comparison of Different Widths****. A. In these plots, width = 0.1, top figure shows estimated output of train data against actual output, and down shows this again but for test. B. In these two figures, width = 200 the rest is the same with A.*

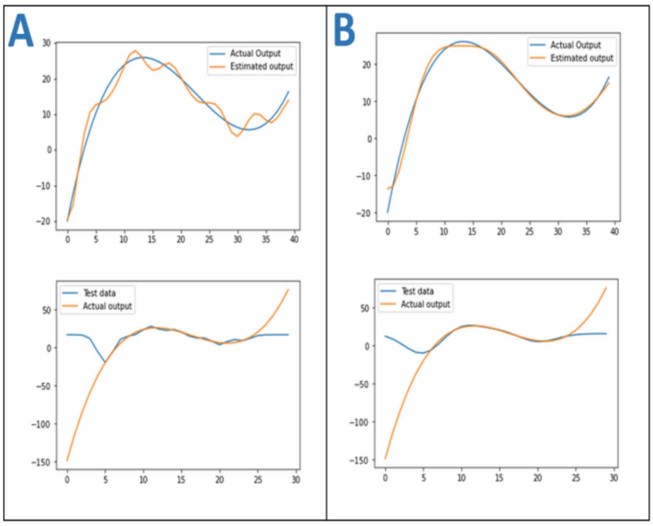

Figure 4 ***Seven Centers with Different Widths.*** *A. top estimated outputs of train data, bottom estimated of test data with variance = 1. B. top approximated output of train data, and bottom, estimated output of test data with variance = 6.*

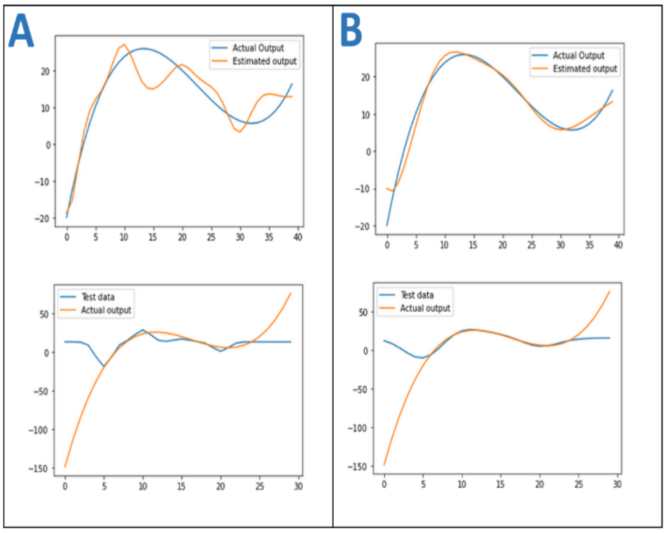

Figure 5 ***Four Centers with Different Widths.*** *A. top estimated outputs of train data, bottom estimated of test data with variance = 1. B. top approximated output of train data, and bottom, estimated output of test data with variance = 6.*

V. References

[1] D. S. Broomhead and D. Lowe, "Multivariable functional interpolation and adaptive networks," 1988.

[2] Broomhead, D. S.; Lowe, David, "Radial basis functions, multi-variable functional interpolation and adaptive networks," 1988.

[3] F. Schwenker, H. A. Kestler and G. Palm, "Three learning phases for radial-basis-function networks," *Neural Networks*, 2001.

[4] Buhmann, Martin Dietrich, "Radial basis functions : theory and implementations.," *Cambridge University Press.*, 2003.