IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 17, Issue 5 Ser. III (Sep. – Oct. 2021), PP 48-51
www.iosrjournals.org

Study and Properties of Orthogonal Neural Network

Dr.Raghvendra Singh
School of Studies in Mathematics, Vikram University, Ujjain, Madhya Pradesh, INDIA

Abstract: *Neural network is an important area of research due to its usage in various fields of engineering and sciences. Orthogonal neural network is a special kind of neural network where basis functions are orthogonal to each other. In this paper, orthogonal neural network is discussed and its important properties are detailed. It is found that its properties very much resemble with Fourier series properties.*

Date of Submission: 16-10-2021 Date of Acceptance: 31-10-2021

I. Introduction

There are three primary elements of significance while making a functional model of the biological neuron. Initially, the synapses of the neuron are displayed as weights. The quality of the association between an input and a neuron is remarked by the estimation of the weight.

Haykin[1] andHaykin et al. [2] coined that values of negative weight show inhibitory connections, while on the other side, positive values assign excitatory associations. An input is provided to the neural network and a corresponding required or target response fixed at the output (in this case, the training is termed as supervised). From the difference between the system output and the desired response, an error occurred [3-5]. This information of error is sustained back to the system and makes adjustments in the system parameters in a proper manner (the learning rule). The procedure continues again and again till the satisfactory results are not accomplished.

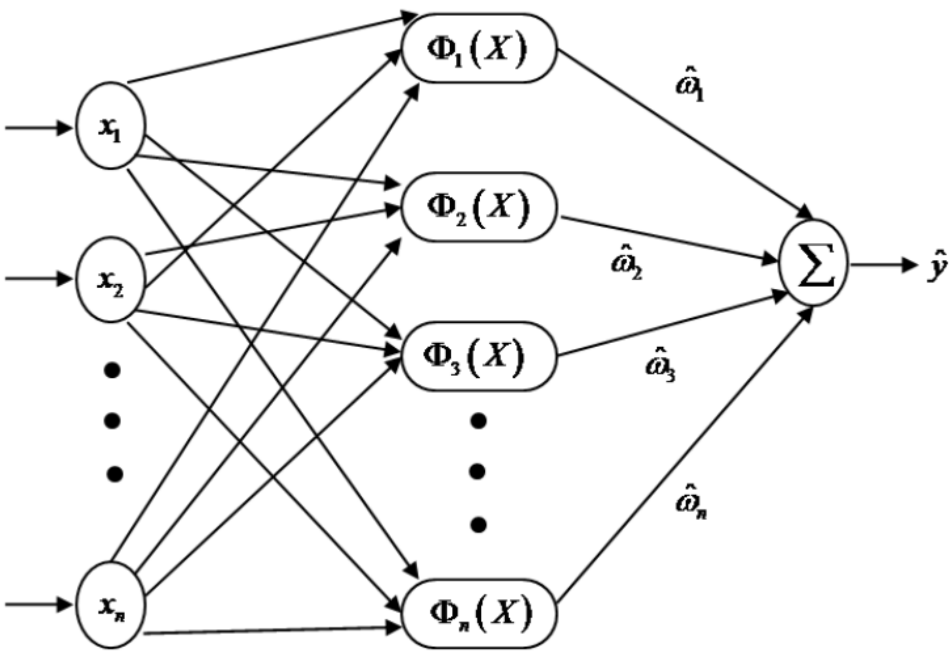

Figure 1: Schematic of orthogonal neural networks

Basic ANN based system suffers from error and slow convergence problems and also suffers from Gibbs phenomenon. To tackle basis ANN problems, Orthogonal Neural Networks (ONN) was proposed. Orthogonal Neural Network is a special kind of feed forward neural network with multiple inputs and single output and a hidden layer with orthogonal activation function in hidden layer.

II. Model of Orthogonal Neural Network

The inputs of neural networks are distributed into orthogonal neurons blocks for each input. The number of neurons for each input signals is arbitrary and i^{th} neuron correspond to i^{th} orthogonal function order [6]. The number of orthogonal neurons is given by

$$N_N = \sum_{i=1}^{m} N_i \qquad (1)$$

where

m – the number of inputs.

N_i– the number of neurons for each input

The other layer of ONN is arranged to nodes which consist of products combinations of the particular outputs from orthogonal neurons and it's defined as

$$\psi_{n_1 \ldots n_m}(t) = \prod_{i=1}^{m} \psi_{n_i}(t_i) \qquad t_i = [t_1, t_2, \ldots, t_m]^T \qquad (2)$$

where

m– dimensions of input vector.

– outputs of orthogonal function implemented by each hidden layer neuron.

The ONN output is given by sum of the all output nodes from previous layer (2) and one is can be mathematically expressed as

$$\hat{\beta}(t,v) = \sum_{n_1}^{N_1} \ldots \sum_{n_m}^{N_m} \omega_{n_1,\ldots,n_m} \psi_{n_1,\ldots,n_m}(t) = \psi^T(t)\hat{\omega} \qquad (3)$$

where

- the transformed input vector.

- the transformed weight vector.

III. Orthogonal Activation Function

The rate of convergence by using orthogonal functions as activation function in neural network is larger than the rate of convergence for perceptron or radial neural networks when it's sticking point. The set of one-dimensional orthogonal functions ψ_{n_i} are defined as [7]:

$$\int_c^d \psi_i^N(t)\psi_j^N(t)dt = \begin{cases} 0 & i \neq j \\ I_i & i = j \end{cases} \qquad (4)$$

If I_i for all i = 1,2, ..., then orthogonal function is orthonormal. Anderson[3] showed that there have been a number of the orthogonal polynomials, e.g. Hermite, Legendre, Laguerre, Chebyshev and Fourier polynomials.

IV. Training Method of Orthogonal Neural Network

Generally, the learning process is performed by adapting the network weight such as the expected value of the mean squared error between network output and training output is minimized. The gradient descent-based learning algorithms are popular training algorithm for neural networks. To train proposed network, learning rules are determined from Lyapunov-like stability analysis. The cost function ONN is given by [8]

$$E_m = \frac{\varepsilon^2}{2} = \frac{1}{2}\left[\beta - \hat{\beta}\right]^2 \qquad (5)$$

where

is learning error, is output of NN, is actual output and is mean square error.

The gradient descent algorithm adjusts network weights in such a way that the square of the neural network learning error changes in a negative gradient direction and then weight adaptation is given by

$$\Delta\widehat{\omega} = -a(t)\frac{\partial E_m}{\partial\widehat{\omega}} \quad (6)$$

where

Δ −Variation of neural network weights,

−Neural network learning rate.

Thus the ONN's weight update law for the instantaneous gradient descent algorithms is given as

$$\widehat{\omega}(t) = \widehat{\omega}(t-1) + a_{\varepsilon}\overline{\Psi}(t)$$

where −Transformed input vector consisting of orthogonal functions.

To guarantee of ONN's training stability, the learning rate α(*t*) must be bounded by [9]

$$0 < a(t) < \frac{2}{\overline{\Psi}\,\overline{\Psi}^T}.$$

V. Properties of ONN

$$\int_c^d \psi_i^N(t)\psi_j^N(t)dt = \begin{cases} 0 & i \neq j \\ I_i & i = j \end{cases} \quad (7)$$

Property 1: If f(x) is any function defined in the interval [c,d] then there exists orthogonal function set $\{\psi_1, \psi_2, \psi_3, \ldots\}$ such that

$$\lim_{N\to\infty}\int_c^d \left(f(x) - \sum_{i=1}^{N}\omega_i\psi_1(x) \right)^2 dx = 0 \quad (8)$$

where, $\omega_i = \int_c^d f(x)\psi_i(x)dx, \quad i = 1, 2, \ldots\ldots$

Property 2: Considering orthogonal function set $\{\psi_1, \psi_2, \psi_3, \ldots\}$ than

$$\begin{aligned} 0 &\le \int_c^d \left(f(x) - \sum_{i=1}^{N}\omega_i\psi_1(x) \right)^2 dx \\ &= \int_c^d f(x)^2 dx - 2\sum_{i=1}^{N}\omega_i\int_a^b f(x)\psi_i(x)dx + \sum_{i=1}^{N}\omega_i^2 = \int_c^d f(x)^2 dx - 2\sum_{i=1}^{N}\omega_i^2 + \sum_{i=1}^{N}\omega_i^2 \\ &= \int_c^d f(x)^2 dx - \sum_{i=1}^{N}\omega_i^2 \\ &\Rightarrow \sum_{i=1}^{N}\omega_i^2 d\int_c^d f(x)^2 dx \end{aligned} \quad (9)$$

Property 3: The approximation error T_n decreases with more number of terms considered.

$$T_N = \int_c^d \left(f(x) - \sum_{i=1}^{N}\omega_i\phi_i(x) \right)^2 dx = \int_c^d f(x)^2 dx - \sum_{i=1}^{N}\omega_i^2$$

$$T_N - T_{N-1} = \int_c^d \left(f(x)^2 dx - \sum_{i=1}^{N}\omega_i^2 \right) - \left(\int_c^d f(x)^2 dx - \sum_{i=1}^{N-1}\omega_i^2 \right) = -\omega_n^2 \le 0 \quad (10)$$

Property 4: A piece-wise continuous function f(x) can be represented by orthogonal set of functions as

$$\sum_{i=1}^{\infty}\omega_i\psi_i(x) = \begin{cases} f(x) & \text{if } f(x) \text{ is continous at } x \\ \frac{1}{2}\left[f(x^-) + f(x^+) \right] & \text{if } (x) \text{ jumps at } x \end{cases} \quad (11)$$

VI. Conclusions

In this paper basic concepts of orthogonal neural network is presneted. The training model with weight updates are discussed. The main four properties of the ONN are detailed and prooof of the properties are presnetd. Th orthogonal set of functions satisfies most of the properites resembing with Foureir series. Property four presnetd above explain Gibbs phenomenon similar to one observed in Fourier series.

References:

[1]. Haykin, S. and Network, N., 2004. A comprehensive foundation. Neural networks, 2(2004), p.41.
[2]. Haykin, S.S., Haykin, S.S., Haykin, S.S. and Haykin, S.S., 2009. Neural networks and learning machines (Vol. 3). Upper Saddle River: Pearson.
[3]. Specht, D.F., 1991. A general regression neural network. IEEE transactions on neural networks, 2(6), pp.568-576.
[4]. Hertz, J., Krogh, A. and Palmer, R.G., 1991. Introduction to the theory of neural computation. Addison-Wesley/Addison Wesley Longman.
[5]. Goldman, M.S., 2009. Memory without feedback in a neural network. Neuron, 61(4), pp.621-634
[6]. Yang, S.S. and Tseng, C.S., 1996. An orthogonal neural network for function approximation. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 26(5), pp.779-785.
[7]. Chen, S., Cowan, C.F. and Grant, P.M., 1991. Orthogonal least squares learning algorithm for radial basis function networks. IEEE Transactions on neural networks, 2(2), pp.302-309.
[8]. White, O.L., Lee, D.D. and Sompolinsky, H., 2004. Short-term memory in orthogonal neural networks. Physical review letters, 92(14), p.148102
[9]. Sher, C.F., Tseng, C.S. and Chen, C.S., 2001. Properties and performance of orthogonal neural network in function approximation. International Journal of intelligent systems, 16(12), pp.1377-1392.

Dr.Raghvendra Singh. "Study and Properties of Orthogonal Neural Network." *IOSR Journal of Mathematics (IOSR-JM)*, 17(5), (2021): pp. 48-51.