A Comparison between Neural Networks and Wavelet Networks in Nonlinear System Identification

Hamed Khodadadi [1], S. Ehsan Razavi [2] and Hossein Ahmadi-Noubari [3]

[1]Department of Technical and Engineering, Khomeini shahr Branch, Islamic Azad University, Isfahan, Iran.
khodadadi@iaukhsh.ac.ir
[2]Department of Electrical Engineering, East Tehran Branch, Islamic Azad University, Tehran, Iran.
erazavi@qdiau.ac.ir
[3]Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada.
noubari@ece.ubc.ca

Abstract: In this paper, identification of a nonlinear function will be presented by neural network and wavelet network methods. Behavior of a nonlinear system can be identified by intelligent methods. Two groups of the most common and at the same time the most effective of neural networks methods are multilayer perceptron and radial basis function that will be used for nonlinear system identification. The selected structure is series - parallel method that after network training by a series of training random data, the output is estimated and the nonlinear function is compared to a sinusoidal input. Then, wavelet network is used for identification and we will use orthogonal least squares (OLS) method for wavelet selection to reduce the volume of calculations and increase the convergence speed.

[Hamed Khodadadi, S. Ehsan Razavi and Hossein Ahmadi-Noubari. **A Comparison between Neural Networks and Wavelet Networks in Nonlinear System Identification.** *J Am Sci* 2018;14(3):42-47]. ISSN 1545-1003 (print); ISSN 2375-7264 (online). http://www.jofamericanscience.org. 6. doi:10.7537/marsjas140318.06.

Keywords: Nonlinear Identification, Nonlinear ARX Model, Radial Basis Function, Wavelet Network, Orthogonal Least Squares

1. Introduction

IN complex systems in which there is vague and inaccurate information, using intelligent methods is common. These methods can predict behavior of system in a non-linear, adaptive way and based on of behavior in human inference method. Among these methods we can refer to artificial neural networks that have been used in various fields including approximation of nonlinear mappings (Harkouss et al, 2011). Artificial neural networks are a kind of vast parallel processor structure that can store experimental information and can make use of them possible. Performance of neural networks is similar to brain from two aspects:

1- Data are gained through training process.

2- Connections between neurons (nerve weights) are used to store information or approximation of nonlinear function.

The processor of neural network is neuron that usually all neurons in a network are similar and can be connected to each other by different methods. Neural network learns information through training process that this information is stored in connection weights between neurons.

For training, the pair of desirable input - output is applied to the network and in each training period, error between actual output and desired output of neural network is used to adjust the weights. By continuing training process, the error is reduced to reach an acceptable level (Pati and Krishnaprasad, 1993; Haykin, 1999). In this paper, two widely used neural networks i.e. multilayer perceptron and radial basis function are used to identify a nonlinear function. Nonlinearity property of neural network is due to nonlinear activation function that is in hidden layer neurons structure and in output. This property makes artificial neural network which can approximate complex nonlinear functions (Harkouss et al, 2011; Haykin, 1999).

In addition to neural networks method we can also benefit from other basic functions to identify and approximate complex nonlinear functions, these methods are known as wavelet transform or wavelet network (wavenet) (Qi and Xiao, 2007; Hung et al, 2004). In neural networks after network structure (number of layers and cells) is determined, initial coefficients of network will be adjusted randomly. Since there is no insight on how each cell operates in network output, applying conventional error minimization methods such as descending gradient usually encounter with problem of involvement in local minimums because this error is highly nonlinear in coefficients (Billings and Wei, 2005; Parasuraman and Elshorbagy, 2005).

According to synthesis analysis relationships, wavelet transform has new possibility for functions approximation or equally identification of a system and based on this extensive efforts have been done to use this potential. In this method, networks with possibility of coefficients adjustment and rapid convergence can be designed by structures similar to neural networks (Billings and Wei, 2005; Srinivas, 2010). Therefore, in following of this paper the mentioned nonlinear function will be also identified by wavelet transform.

2. Problem Explanation

Identification and characterization are basic issues in control theory which are used to explain a system mathematically. In general, the model of a dynamic system can be shown by following relationships:

$$dx(t)/dt = \dot{x}(t) = \Phi[x(t), u(t)] \qquad t \in R^+$$
$$y(t) = \Psi[x(t)]$$
(1)

In above relationship Φ and Ψ are static nonlinear functions that explain relationships of state variables, inputs and outputs. Also x is system state vector, u is system inputs vector and y is system output vector (Khalil, 2002).

$$x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T \qquad (2)$$

$$u(t) = [u_1(t), u_2(t), \dots, u_p(t)]^T \qquad (3)$$

$$y(t) = [y_1(t), y_2(t), \dots, y_m(t)]^T \qquad (4)$$

These relationships can also be expressed in discrete time in which variable t is replaced by variable k.

$$x(k+1) = \Phi[x(k), u(k)] \qquad t \in R^+$$
$$y(k) = \Psi[x(k)] \qquad (5)$$

2.1 Identification Models of Linear System

For linear time-invariant systems with unknown parameters, making identification models is a known issue. Ljung (1999) For a SISO system proposed a controllable and observable model that is expressed by relationship (6).

$$y_P(k+1) = \sum_{i=0}^{n-1} \alpha_i . y_P(k-i) + \sum_{j=0}^{m-1} \beta_j . u(k-j) \qquad (6)$$

where α and β are unknown coefficients. This relationship shows that the value of system at time $k+1$ is equal with linear combination of past moments values in function and input. For doing identification process two following methods called parallel and series - parallel can be used (Haykin, 1999).

Parallel method:

$$\hat{y}_P(k+1) = \sum_{i=0}^{n-1} \hat{\alpha}_i(k).\hat{y}_P(k-i)$$
$$+ \sum_{j=0}^{m-1} \hat{\beta}_j(k).u(k-j) \qquad (7)$$

Series - parallel method:

$$\hat{y}_P(k+1) = \sum_{i=0}^{n-1} \hat{\alpha}_i(k).y_P(k-i)$$
$$+ \sum_{j=0}^{m-1} \hat{\beta}_j(k).u(k-j) \qquad (8)$$

The difference between two equations is y entered into the identifier system. In one equation y is obtained from the main system and in other equation y is identified from system. For making stable adjustment rules using series - parallel method is preferred. In this paper the selection method is also series - parallel method.

2.2 Identification Models of Nonlinear System

Various models can be used to identify a nonlinear function. Model used in this paper has the following form.

$$y_P(k+1) = f[y_p(k), y_p(k-1), \dots,$$
$$y_p(k-n+1), u(k), u(k-1), \dots, u(k-m+1) \qquad (9)$$

In these relationships u and y are input and output of system, respectively. Here similar to linear state there are two types of parallel and series - parallel methods that the second method is selected.

Block diagram of this model and series - parallel method is indicated in figure 1 and figure 2 respectively. These diagrams show that what is input to neural network, what is output from it and there are how many blocks of memory (due to delays).

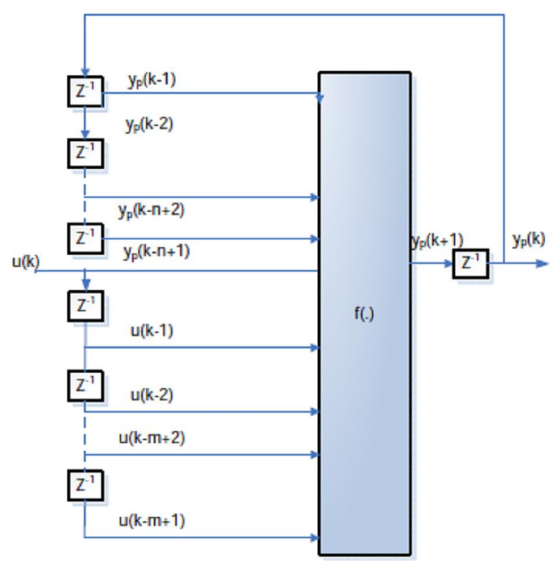

Figure 1 - Diagram of used structure for nonlinear system identification

For wide range of nonlinear systems, the nonlinear autoregressive with exogenous inputs (NARX) model can be considered as a representer

(Billings and Wei, 2005). For example a SISO system like (10) will be evaluated.

$$y_P(k+1) = f[y_P(k), y_P(k-1)] + u(k)$$
$$= \frac{y_P(k).y_P(k-1).[y_P(k)+2.5]}{1+y_P^{\,2}(k)+y_P^{\,2}(k-1)} + u(k) \qquad (10)$$

For network training a random input in the interval [-1,1] is applied to systems. After determining the weights of neural network, system output to specified sinusoidal input will be predicted based on series - parallel method.

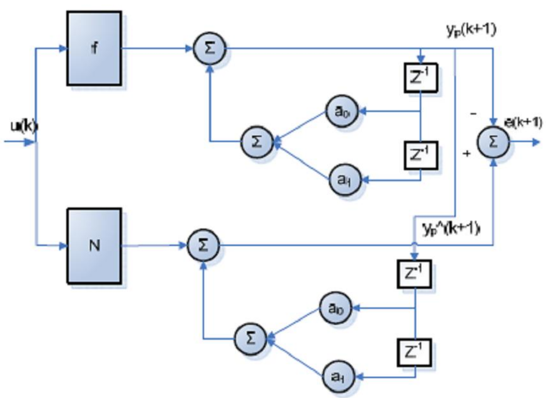

Figure 2 - Block diagram of series – parallel identification

3. Multilayer Perceptron Neural Network

Perceptron neural network is an almost all-purpose neural network which is used in most applications and the obtained results of it have always been acceptable. This network with backpropagation error algorithm has been introduced in detail in different books like Haykin (1999). Here briefly the mathematical relationships governing on the network are given.

$$y = \rho\left(\sum_j w_j^O Q_j + b^O\right), \; Q_j = \rho(q_j)$$
$$q_j = \sum_i w_{ij}^I u_i + b_j^I \;, \; \rho(\cdot) = \frac{1}{1+e^{-(\cdot)}} \qquad (11)$$

where w_{ij}^I and w_j^O are input and output weights, b_j^I and b^O are hidden layer bias and output weights respectively and u_i is neural network input. The structure of this neural network is shown in Figure 3.

Backpropagation error algorithm can be used for training of this network. We can make the network behavior near to desirable state by using mentioned algorithm and data from system model.

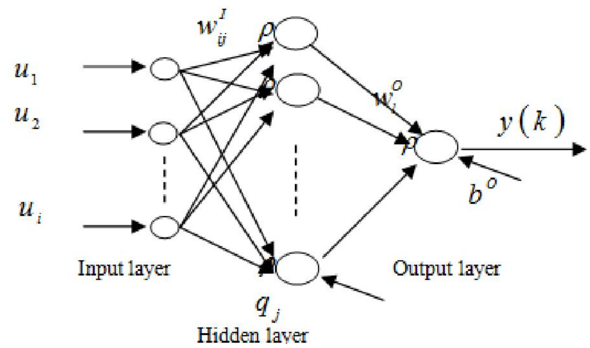

Figure 3 – Structure of multilayer perceptron neural network

In this part MLP network is used to identify the dynamics of introduced nonlinear function. System inputs in this case are function value at any moment and a moment before it and also neural network output is prediction of function value in next moment. By applying a random input in interval [-1,1] to the trained network, final error between estimated output and output of main system and response of both system to sinusoidal input is compared in figure 4 and figure 5 respectively.

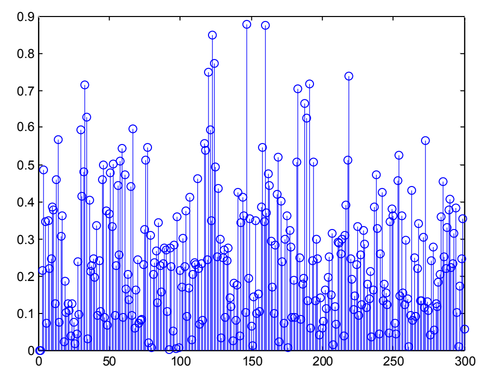

Figure 4 – Identification error for MLP method

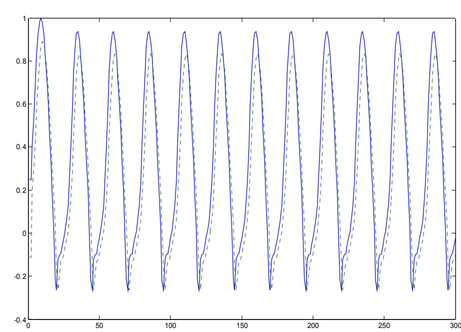

Figure 5 - Comparing identified system output (dashed line) and main system output (filled line) in MLP method

4. RBF Neural Network

Neural networks are capable of learning and can adapt themselves with environment. Because of this capability, when mathematical model of system is not available or complex calculations are required for reaching to this model, neural networks can be very useful (Pati and Krishnaprasad, 1993). RBF network shown in Figure 6 which consists of three layers including input layer, hidden layer and output layer can be considered as a mapping from space $\Re^r \rightarrow \Re^s$.

$P \in \Re^r$ is considered as input vector and $C_i \in \Re^r \quad (1 \le i \le u)$ as centers of hidden units are considered. The output of each RBF unit is calculated as follows.

$$R_i(P) = R_i(\|P - C_i\|) \qquad i = 1, \ldots, n \qquad (12)$$

So that $\| \ \|$ indicates Euclidean distance in the input space. Usually, Gaussian functions are used for RBF units.

$$R_i(P) = \exp\left[-\frac{\|P - C_i\|^2}{\sigma_i^2}\right] \qquad (13)$$

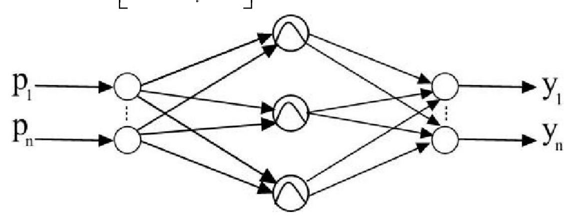

Figure 6 - Structure of RBF neural network

where σ^i shows unit width of i^{th} RBF. $Y_j(P)$, (j^{th} output) in RBF neural network is calculated as follows.

$$Y_i(P) = \sum_{i=1}^{n} R_i(P) \times W(j,i) \qquad (14)$$

$W(j,i)$ is weight between i^{th} RBF unit and j^{th} output. According to above relationships it can be seen that output of RBF classifiers is characterized by linearly separable functions. They develop decision boundaries in output space. Capability of RBF neural networks is strongly dependent on separability of classes in u-dimensional space that is done by RBF units and by a non-linear transformation. Geometrically, the idea of RBF neural network is separation of input space in several sub-spaces as spherical regions (circle) (Haykin, 1999; Qi and Xiao, 2007).

Also, in this section one RBF network is used for modeling a nonlinear function. The training method is still series - parallel. In effect, input to neural network is output of main system with its delayed version. Training error, main system response and estimate of RBF network are shown in figure 7 and figure 8.

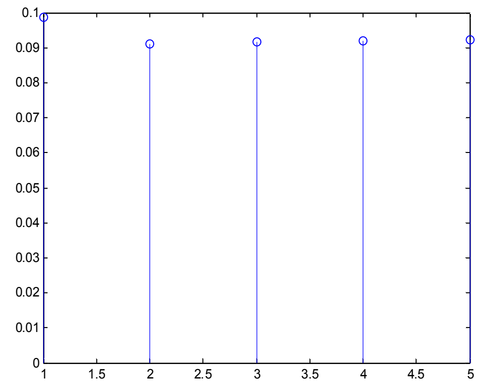

Figure 7 - Training error during RBF method

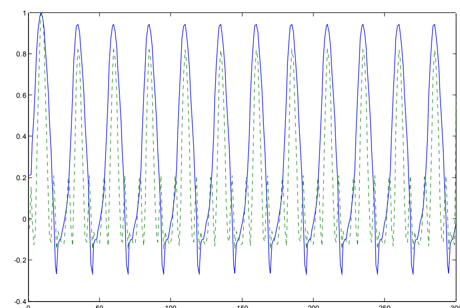

Figure 8 - Comparing identified system output (dashed line) and main system output (filled line) in RBF method

5. Wavelet Transform

The main idea of wavelet transform is to depict functions from $L^2(R)$ space to time-frequency so that transform coefficients can express characteristics of time-frequency signal. Continuous wavelet transform gives a picture of function in developed space by shift and scale of a special function called mother wavelet. When $\psi(x)$ is mother wavelet some sets of functions are applied for continuous wavelet transform as follows (Parasuraman and Elshorbagy, 2005; Ismail et al, 2009).

$$h_s(a,b) = |a|^{-\frac{1}{2}} . h_s\left(\frac{x-b}{a}\right) \qquad a, b \in R, a \neq 0 \qquad (15)$$

and we have:

$$f(x) = \frac{1}{C_4} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} < h_s(a,b), \quad f > h_s(a,b)\frac{da}{a^2} db \qquad a \neq 0 \qquad (16)$$

$$C_4 = \int_{-\infty}^{+\infty} \frac{|h_s(w)|}{|w|} dw < \infty \qquad (17)$$

where ψ is a limit set of Z^2. In above relationship *f (X)* is approximated regarding elements of the set. The above relationship compared with neural networks is called wavelet network. This relationship identifies the network for approximating one variant relationship. To extend this network to a form of multivariant network it seems using products forms is the easiest method, i.e.:

$$\psi(x_1,...,x_N) = \varphi\left(\frac{x_1 - t_1}{d_1}\right).....\varphi\left(\frac{x_n - t}{d_n}\right) \qquad (18)$$

The problem of this method is large volume of computations that develops because of applying a nonlinear function ψ. The solution which is proposed by Hung et al (2004) for this problem is use of radial wavelets i.e. wavelet that satisfies in the following equation.

$$\psi(X) = \psi(\|X\|) \qquad (19)$$

where $\|X\| = (X^T X)^{1/2}$ and ψ is a single-input-single-output function. Therefore, the volume of performed calculations in every cell becomes independent of dimensions of variables and only is limited to a soft calculation and applying nonlinear function ψ. By using radial wavelet the network form in multivariant state becomes as follows:

$$g(x)\sum_{i=1}^{N} w_i \varphi(diag(di(x - ti))) \qquad (20)$$

where N is the number of cells used and parameters $d_i \in R^n, L_i \in R^n, w_i \in R$ are weights vector, transfer matrix and change of scale, respectively. It will be proved that the above set can approximate every function in $L^2(R^n)$.

Hence, the wavelets' family can be considered as follows:

$$h_{m,k}(x) = a_0^{-\frac{m}{2}} \varphi(a_0^{-m} x - kb_0) \quad m,k \in z \qquad (21)$$

where ψ is the mother wavelet and $a,b \in Z, X \in R^n$.

In this section to identify nonlinear function, Mexican hat wavelet with following relationship is used for mother wavelet:

$$\varphi(x) = (2/\sqrt{3}\pi^{-1/4})(1 - x^2)e^{-x^2/2} \quad m,k \in z \qquad (22)$$

Diagram of this wavelet is according to the figure 9.

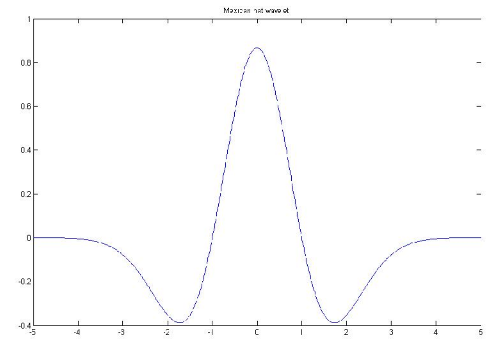

Figure 9 - Diagram of Mexican hat mother wavelet

We want to apply limited number of resulting set of wavelets for nonlinear function approximation. For this purpose we must select those that by least number approximate the nonlinear function in the best way. This work can be done by orthogonal least squares (OLS) method and introduced by Chen et al (1989). Modeling result of nonlinear function using Mexican hat wavelet is shown in Figure 10.

6. Conclusion

In this paper, neural network and wavelet network methods are used to identify a nonlinear function. Two first reviewed methods use series - parallel method. In these methods at first neural network was trained by random data and then sinusoidal data was given to the system. During applying sinusoidal input, future output was predicted by delayed outputs. Two methods MLP and RBF operation in identification of nonlinear system is almost identical, although MLP has a higher speed. Then, wavelet method was evaluated and it was expressed that although the use of a regression model developed based on wavelet expansion is useful to identify low-dimensional systems, after increase of dimension, this method will confront with some problems.

In order to solve this problem the wavelet network was used. Obviously, for escaping from problems and having advantages of wavelet and regression simultaneously, radial wavelet network was reviewed and it was shown that by using a radial mother wavelet rather than separable one, a model similar to RBF neural networks can be achieved. Also, using some linear regressors along with non-linear regressors (wavelet) will accelerate the identification. Hence, the selected wavelet network method showed better performance compared with other methods from accuracy and fitting aspects and also from the aspect of speed of convergence.

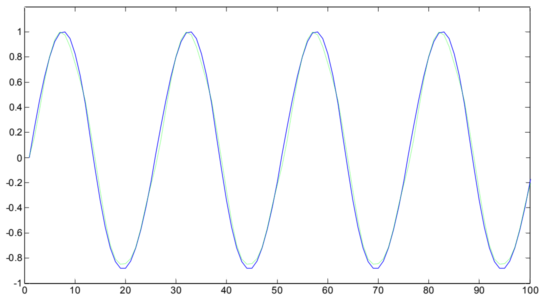

Figure 10 - Comparing identified system output (dashed line) and main system output (filled line) in wavenet method

Corresponding Author:
Hamed Khodadadi
Department of Technical and Engineering,
Khomeini shahr Branch,
Islamic Azad University,
Isfahan, Iran.
Email: khodadadi@iaukhsh.ac.ir

References

1. Billings, S.A. and Wei, H.L., 2005. A New Class of Wavelet Networks for Nonlinear System Identification. IEEE Transaction on Neural Networks, Vol. 16, No. 4: 862-874.
2. Chen, S., Billings, S.A. and Luo, W., 1989. Orthogonal Least Squares Methods and their Application to Non-linear System Identification. International Journal of Control, Vol. 50, No. 5: 1873-1896.
3. Harkouss, Y., Fahs, W., and Ayache, M., 2011. A New Algorithm for Structure Optimization of Wavelet Neural Network. International Journal of Computer Science Issues, Vol. 8, Issue 2: 108-117.
4. Haykin, S., 1999. Neural Networks, a Comprehensive Foundation. 2nd Edition, Prentice Hall, New Jersey.
5. Hung, S.L., Huang, C.S. and Wen, C.M., 2004. Using Wavelet Neural Network for the Identification of A Building Structure from Experimental Data. 13th World Conference on Earthquake Engineering, Vancouver, B.C., Canada, August 1-6, Paper No. 241.
6. Ismail, N., Tajjudin, N., Rahiman, M.H.F. and Taib, M.N., 2009. Estimation of Nonlinear ARX Model for Steam Distillation Process by Wavenet Estimator. Proceeding of 2009 IEEE Student Conference on Research and Development, UPM Serdang, Malaysia, Nov. 16-18: 562-565.
7. Khalil, H.K., 2002. Nonlinear Systems. 3rd Edition, Prentice Hall, New Jersey.
8. Ljung, L., 1999. System Identification: Theory for the User. 2nd Edition, Prentice Hall, New Jersey.
9. Parasuraman, K. and Elshorbagy, A., 2005. Wavelet Networks: An Alternative to Classical Neural Networks. Proceedings of International Joint Conference on Neural Networks, Montreal, Canada, July 31 - August 4: 2674- 2679.
10. Pati, Y.C. and Krishnaprasad, P.S., 1993. Analysis and Synthesis of Feedforward Neural Networks using Discrete Affine Wavelet Transformations. IEEE Transaction on Neural Networks. Vol. 4, No. 1: 73-85.
11. Qi, W. and Xiao, S., 2007. SINS/GPS Integrated Navigation for Autonomous Underwater Vehicle with Wavelet Package Analysis and Neural Networks. 8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing. Qingdao, China. July 30 - August 1: 1028-1033.
12. Srinivas, H.K., 2010. Application of Artificial Neural Network and Wavelet Transform for Vibration Analysis of Combined Faults of Unbalances and Shaft Bow. Advances in Theoretical and Applied Mechanics, Vol. 3, No. 4: 159 – 176.

3/17/2018