SYSTEM IDENTIFICATION USING NEURAL NETWORKS

Abid Ali
Fakultät für Elektrotechnik und Informationstechnik, Ruhr-Universität Bochum, Germany

Christian Schmid
Institute of Automation and Computer Control, Ruhr-Universität Bochum, Germany

Keywords: artificial neuron, neural network, learning algorithm, identification, model, black-box modeling, feed-forward neural network, multi-layer perceptrons, radial-basis function networks, local model networks, recurrent networks, learning, back-propagation, back-propagation through time, continuous-time systems, discrete-time systems, structure optimization, parameter estimation

Contents

Summary

This chapter describes how artificial neural networks (ANN) can be used for the identification of nonlinear dynamic systems. Some static artificial neural networks like multi-layer perceptrons, radial-basis function networks and local model networks are treated briefly. A short description of some continuous as well as discrete-time dynamic networks is also given.

Training techniques associated with each network for the optimization of structure of the network and parameter estimation are also listed. Different identification approaches based on static as well as dynamic networks reported in the literature dealing with the identification of discrete and continuous-time systems using input-output as well as state-space models are summarized. Some practical issues related to neural identification are also discussed. A bibliography is included for an in-depth study of the

subjects presented here.

1. Introduction

Research in the field of artificial neural networks (ANN) is inspired by the biological nervous systems. Artificial neural networks are composed of simple elements, known as artificial neurons, operating in parallel. As in nature, the network function is determined largely by the connections between these elements. These connections are known as synapses. A neural network can be trained to perform a particular function by adjusting the synaptic weights between its elements.

Commonly, synaptic weights of neural networks are adjusted, or trained, so that a particular input leads to a specific target output. Mostly, a network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically many such input-target pairs are used to train a network.

Neural networks are characterized by their inherent nonlinearity, capability to learn, universal approximation property, parallel processing and modular structure. Typical applications of neural networks include pattern recognition, classification, function approximation, system identification, speech recognition, vision and control.

In late nineteen eighties the research interest in the capabilities of artificial neural networks as approximators of arbitrary continuous functions was put into focus. This was the time when artificial neural networks found another interesting area of application, the nonlinear system identification (see *Identification of Nonlinear Systems*). This interest reached its climax in mid 1990s. The earlier concepts used multi-layer perceptron (MLP) networks for identification tasks. In the meantime researchers began to use radial-basis functions (RBF) to construct neural networks.

Due to the linear-in-parameter property of the output layer, RBF networks attracted the attention of people working on approximation and system identification. MLP and RBF networks have proven themselves good means for nonlinear system identification. The major drawback of such black-box models is that the network parameters have no direct correspondence to the physical system parameters.

If the parameters of a linearized model of the system at current point of operation or the gradient of some objective function are to be calculated, then the gradient of this neural model has to be calculated. This calculation is computationally expensive.

Moreover, the presence of over- or under-fitting ripples in RBF or MLP network-based approximations may lead to erroneous gradient calculations. In the past ten years much attention has been given to the application of another type of networks, known as local model networks (LMN), to system identification.

The most popular form of these networks uses local linear models (LLM). The local model networks decompose the input space of the nonlinear mapping into different local regions and estimate a linear model for each region.

If the number of regions is sufficiently large and the operating regions overlap properly then any nonlinear mapping can be approximated smoothly. This smoothness of approximation is necessary if a model-based controller is to be designed, or the model gradients are required for some optimization purposes.

Another major advantage of such networks is the exploitation of linear design techniques while devising corresponding controller networks. One can find in the literature some identification approaches based on other types of neural networks like Cerebellar Model Arithmetic Computer (CMAC) and B-spline networks.

All the networks described above do not contain internal dynamics. In order to use these networks for the identification of nonlinear dynamic systems external dynamic elements are necessary. The networks having internal dynamics like dynamic multi-layer perceptrons and recurrent networks, have also been tried for identification purposes. But due to their internal dynamics, stability problems must be considered.

Training of these networks normally possesses relatively poor convergence properties when compared with static networks. Research in the field of neural identification has been focused mainly on estimating discrete-time nonlinear input-output models. Such models are described by nonlinear difference equations.

A little attention was given to the identification of nonlinear state-space models. Neural identification of nonlinear continuous-time models is also under-represented in the literature. See *General Models of Dynamic Systems* for description of these models.

This chapter describes some application possibilities of neural networks in nonlinear system identification. As there is no chapter in the whole theme dedicated to the introduction of artificial neural networks, the next section is devoted for a brief introduction of some network paradigms and their learning algorithms. Static as well as dynamic networks are introduced in this section.

As the objective is system identification, the neural paradigms are described from the approximation point of view. Section 3 describes how these neural networks can be applied to the identification of nonlinear dynamic systems. Different approaches for the identification of discrete as well as continuous-time models are presented.

2. Artificial Neural Networks

A neuron is the smallest information processing unit in a neural network. A simple model of an artificial neuron is given in Figure 1. Variables $u_1, u_2, ..., u_r$ stand for the inputs to the neuron and y represents its output. All the inputs multiplied by their respective synaptic weights $w_1, w_2, ..., w_r$ are summed up in the summing junction.

A bias term b is subtracted from this weighted sum to determine the internal activity a of the neuron. The output y is a nonlinear function of the internal activity and can be given as

$$y = f(a) = f(\sum_{i=1}^{r} w_i u_i - b) \tag{1}$$

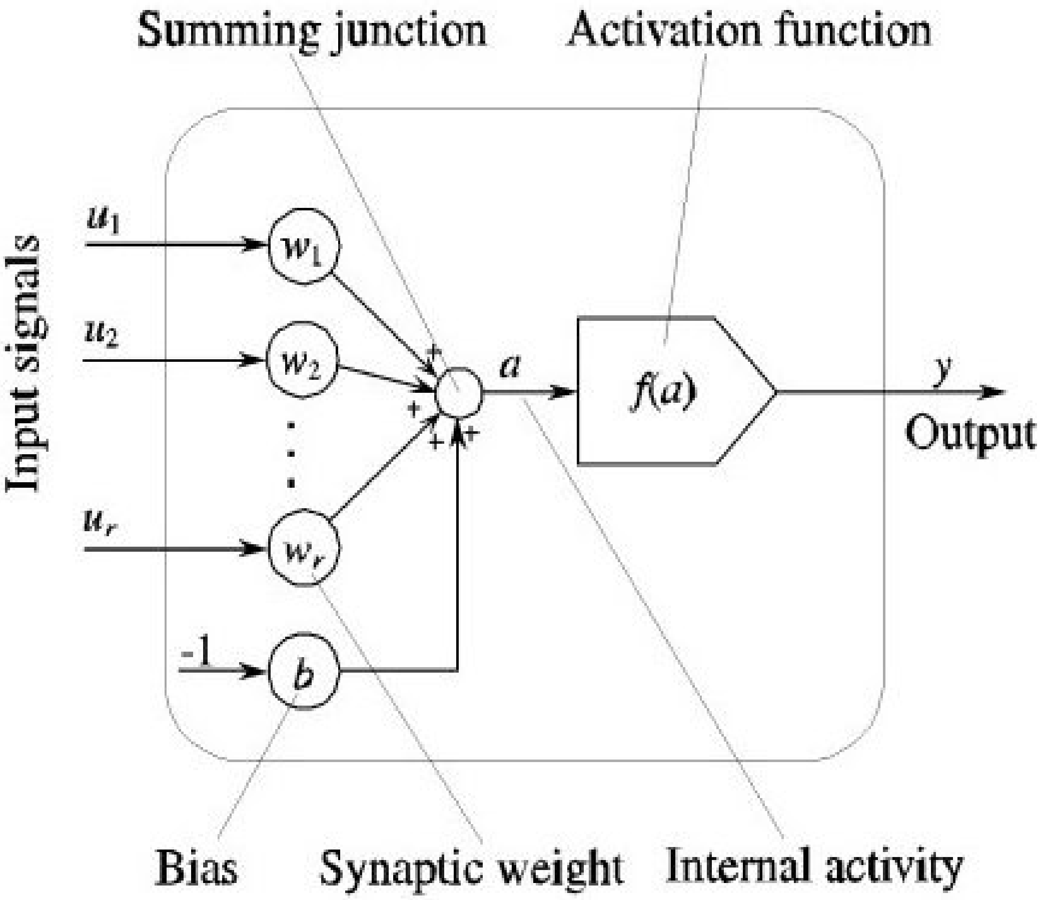

Figure 1: A simple model of an artificial neuron

The function f is known as the activation or squashing function, which maps the internal activity a to a closed interval $[0,1]$ or alternatively $[-1,1]$. There are several types of this function reported in the literature, e.g. threshold functions, saturated linear functions and sigmoid functions.

These functions are described in Table 1. Neurons with threshold functions are binary processing units and are useful for classification and decision tasks. For the identification purposes, where a smooth approximation of a nonlinear function is desired, sigmoid functions are preferred.

Another advantage of the sigmoid functions is that these functions are differentiable, which is a requirement when gradient-based learning techniques are to be applied to train the neural network.

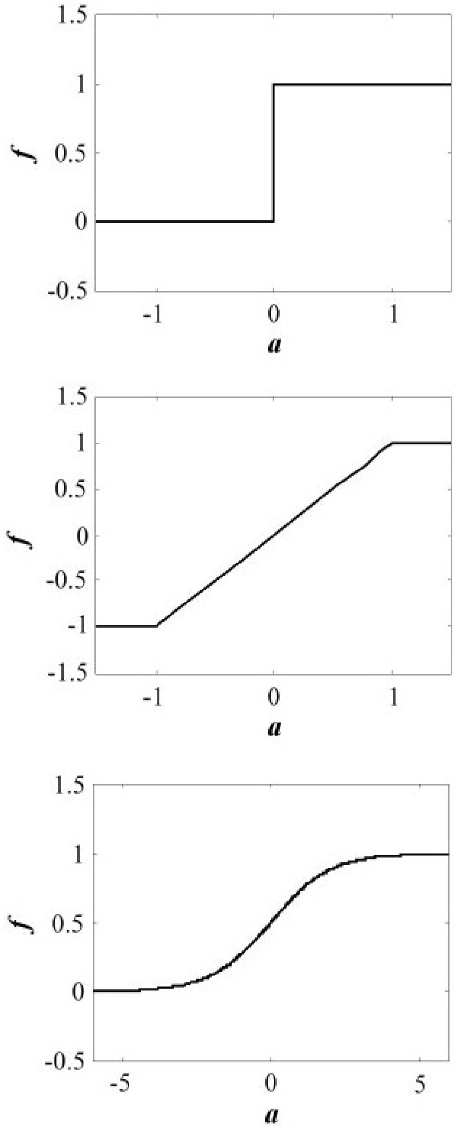

Table 1: Some examples of activation functions

A neural network is composed of a small or large number of neurons which are coupled with each other. In general these neurons are arranged in different layers. Neural networks consisting of more than one layer are termed as multi-layer networks.

Based on the different possibilities of neuron interconnections the networks can be categorized as static or dynamic networks.

Before a network can perform the desired task, it should be trained. Training is performed by feeding the network with a set of data patterns and adjusting its synaptic weights in order to achieve a desired response to input data. The algorithms used to train the networks are known as learning algorithms.

A learning algorithm is a set of rules, which are applied during the training phase to adjust the parameters of the neural network in order to perform better.

-
-
-

Bibliography

Bohn C. (2000). *Recursive Parameter Estimation for Nonlinear Continuous-Time Systems through Sensitivity-Model-Based Adaptive Filters*. Ph.D. thesis, Ruhr-Universität, Bochum. [A good treatment of sensitivity-model-based methods for the identification of nonlinear continuous time systems]

Brown M., Bossley K.M., Harris C.J. (1996). The theory and implementation of the B-spline neurofuzzy construction algorithms. In *Proceedings of the Fourth European Congress on Intelligent Techniques and Soft Computing (EUFIT '96) September 1996*, vol. 2, pp. 762 – 766, Aachen, Germany. [This paper describes some of the design choices for the development of an adaptive neurofuzzy system based on B-splines]

Funahashi K. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks* 2, 183 – 192. [Proves that any continuous mapping can be approximately realized by multi-layer neural networks with at least one hidden layer].

Funahashi K., Nakamura Y. (1993). Approximation of dynamical systems by continuous time recurrent networks. *Neural Networks* 6, 801 – 806. [This paper deals with the trajectory approximation by continuous-time recurrent neural network]

Gupta M.M., Rao D.H. (1994). *Neuro-Control Systems*. Piscataway, NJ: IEEE Press. [This book brings together a number of historical papers important to the practical application of neural networks to identification and control]

Haykin S. (1994). *Neural Networks - A Comprehensive Foundation*. New York: Macmillan College Publishing Company. [A comprehensive treatment of fundamentals of neural networks]

Himmelblau D. (1972). *Applied Nonlinear Programming*. New York: McGraw-Hill Book Company. [This book deals with optimization techniques]

Hunt K., Johansen A. (1997). Design and analysis of gain-scheduled control using local controller networks. *International Journal of Control* 66, 619–651. [This papers deals with application of local model networks to design gain-scheduled controllers]

Hunt K.J., Irwin G.R., Warwick K. (1996). *Neural Network Engineering in Dynamic Control Systems*. London: Springer-Verlag. [A collection of different papers on applications of neural networks to control and identification]

Hunt K.J., Sbarbaro D., Żbikowski R., Gawthrop P.J. (1992). Neural networks for control systems - a survey. *International Journal of Control* 28, 1083 – 1112. [A survey of pioneering approaches of the neural identification and control]

Jin L., Nikiforuk P.N., Gupta M.M. (1995). Approximation of discrete-time state-space trajectories using dynamic recurrent neural networks. *IEEE Transactions on Automatic Control*, 40, 1266 – 1270. [Deals with the approximation capability of a class of discrete-time dynamic recurrent neural networks]

Johansson R. (1993). *System Modelling & Identification*. Englewood Cliffs, New Jersey: Prentice-Hall International, Inc. [A good book on system identification. Parameter estimation techniques for discrete as well as continuous-time systems are presented]

Junge T. (1999). *On-line Identifikation und lernende Regelung nichtlinearer Regelstrecken mittels neuronaler Netze*. Ph.D. thesis, Ruhr-Universität, Bochum. [This thesis deals with application of RBF and local models networks to identification and control based on input-output models].

Leistritz L., Galicki M., Witte H., Kochs E. (2002). Training trajectories by continuous recurrent multilayer networks. *IEEE Transactions on neural networks* 13, 283 – 291. [A treatment of continuous-time dynamic multi-layer neural networks]

Ljung L., S¨oderstr¨om T. (1983). *Theory and Practice of Recursive Identification*. Cambridge: MIT Press. [A very good and general treatment of prediction error based recursive identification methods]

McLoone S.C., Irwin G., McLoone S.F. (2001). Constructing networks of continuous-time velocity-based models. *IEE Proceedings Control Theory and Applications* 148(5), 397 – 405. [This paper deals with estimation of velocity-based local models for continuous-time systems]

Miller W.T., Glanz F.H., Kraft L.G. (1990). CMAC: An associative neural network alternative to back-propagation. *Proceedings of the IEEE* 78, 1561 – 1567. [This paper describes the cerebellar model arithmetic computer (CMAC) neural network as an alternative to back-propagated multilayer networks]

Murray-Smith R., Johansen T.A. (1997). *Mutliple Model Approaches to Modelling and Control*. London: Taylor & Francis Ltd. [This book brings together different multiple model approaches like local model networks and fuzzy systems etc]

Narendra K., Parthasarathy K. (1990). Identification and control of dynamical systems using neural networks. *IEEE Transactions on neural networks* 1, 4–27. [One of the pioneering works on the modeling and identification and control of nonlinear systems using neural networks]

Poggio T., Girosi F. (1990). Networks for approximation and learning. *Proceedings of the IEEE*

78, 1481 – 1497. [A major contribution to the theory, design and applications of radial basis function networks based on regularization theory]

Qin S.Z., Hu H.T., McAvoy T.J. (1992). Comparison of four neural net learning methods for dynamic system identification. *IEEE Transactions on neural networks* 3, 122 – 130. [A contribution on the comparison of different learning methods for feedforward and recurrent neural networks]

Rojas R. (1996). *Neural Networks - A Systematic Introduction*. Berlin: Springer-Verlag. [A good text book on fundamentals of neural networks]

Sjöberg J. (1995). *Non-linear System Identification with Neural Networks*. Ph.D. thesis, University

of Linköping, Sweden. [This thesis gives a unified overview of nonlinear black-box modeling and investigates the use of neural networks in system identification]

Unbehauen H. (1995). *Regelungstechnik* III. Braunschweig: Vieweg-Verlag, 5 edn. [A good German text book which also deals with identification and parameter estimation]

Unbehauen H., Rao G.P. (1990). Continuous-time approaches to system identification – a survey. *Automatica* 26(1), 23–35. [A survey of different identification schemes for continuous-time systems].

Young P. (1981). Parameter estimation for continuous-time models – a survey. *Automatica* 17(1), 23–39. [A survey paper on different parameter estimation approaches for continuous time models]

Żbikowski R., Hunt K.J. (1996). *Neural Adaptive Control Technology*. Singapore: World Scientific Publishing. [A collection of different papers on adaptive control and identification using neural networks]

Biographical Sketches

Abid Ali was born in Haroonabad, Pakistan. He obtained his Bachelor's Degree in Electrical Engineering from University of Engineering and Technology Lahore, Pakistan, in 1991. During 1991-1993, he worked for Siemens (Pakistan) and other companies in the field of protection, instrumentation and control of power systems and power house equipment. In 1998, he received his Diplom in Electrical Engineering from Ruhr-Universität Bochum, Germany. Currently, he is working at the Institute of Automation and Computer Control, Ruhr-Universitäty Bochum. His current research interests are in nonlinear and adaptive control, neural networks, local model networks and application of multimedia in control engineering education.

Christian Schmid received his Diploma Degree in Mechanical Engineering in 1972 from the University of Stuttgart, and a Dr.-Ing. Degree in Electrical Engineering from Ruhr-Universität Bochum in 1979. From 1972 to 1975, he was Research Assistant at the University of Stuttgart. From 1975 to 1992, he was assistant professor and senior lecturer in the Control Engineering Laboratory of Ruhr-Universität Bochum. In 1992 he made the habilitation in Control Engineering. In 1993 he was awarded the title of Privatdozent and in 2000 he was promoted as Professor of Control Engineering in the Department of Electrical Engineering and Information Sciences at Ruhr-Universität Bochum. His main research interests were numerical methods in control, model reduction, optimization and are currently in the field of computer-aided control system design, adaptive control, applications using neural networks and modeling. He has started projects in the area of multimedia-based learning, virtual and remote labs with German and European partners and is currently managing a nation-wide project on networked learning in the area of control education.