

Chapter III

Kolmogorov's Spline Complex Network and Adaptive Dynamic Modeling of Data

Boris Igelnik
BMI Research Inc., USA

ABSTRACT

This chapter describes the clustering ensemble method and the Kolmogorov's Spline Complex Network, in the context of adaptive dynamic modeling of time-variant multidimensional data. The purpose of the chapter is to give an introduction in these subjects and to stimulate a participation of both young and experienced researchers in a solution of challenging and important for theory and practice problems related to this area.

INTRODUCTION

This chapter describes specific neural network architecture with complex weights and potentially complex inputs in the context of adaptive dynamic modeling of *time-varying* multidimensional data. The technological and scientific developments in many areas of human activity have reached a level, requiring adequate changes in traditional methods of data modeling. Consider just two examples. One is related to coal-operating power stations. Coal might provide a fuel for world power industry for hundreds of years, making a good alternative to rapidly decreasing oil. But coal combustion produces harmful pollutions. Mitigation of this problem by controlling combustion requires modeling of power data, which are time-variant, highly multidimensional, nonlinear, non-stationary, and influenced by complicated, interacting chemical, electro-magnetic, and mechanical processes. There is no way to do such modeling by traditional methods of control theory (Ljung, 2000; Astrom & Wittenmark, 1995). The second example is related to defense, in particular to problems of detection, identification, and tracking targets in the clutter environment, utilizing sensors, such as radar, sonar, infrared (Hovanessian, 1984; Scolnik, 1990), and others. A possibility of having multiple moving and interacting targets, clutters, and sensors makes these problems extremely difficult for solution in real applications. The methods for solution of these problems, basically Bayesian ones (Antony, 1995; Congdon, 2006; Stone, Barlow, & Corvin, 1999), are founded on the theory developed 30-50 years ago, and inadequate to currently existing reality.

There is a long history of signal and noise representation, utilizing complex numbers in signal processing (Oppenheim & Schaffer, 1975; Rihaczek & Hershkowitz, 1996; Haykin, 2001). Relatively recently it was recognized that complex representation of inputs and adaptively adjusted weights may be helpful in *neural network (net)* modeling, especially for pattern recognition (Kim & Guest, 1990; Leung & Haykin, 1991; Georgiu & Koutsougeras, 1992; Nitta, 1997, 2003; Arena, Fortuna, Muscato, & Xibilia, 1998; Aizenberg, Aizenberg, & Vandewalle, 2000; Igelnik, Tabib-Azar, & Leclair, 2001a; Hirose, 2003, 2006).

This chapter has the following objectives: 1) introducing basic principles, ideas, and algorithms of adaptive neural net modeling of *time-varying*, highly multidimensional data; 2) introducing a specific complex-valued neural network, the *Kolmogorov's Spline Complex Network (KSCN)*, which might be advantageous especially in various tasks of pattern recognition.

BACKGROUND

There are several approaches to modeling, divided into two intersecting groups of methods, Artificial Intelligence (AI) (Nilsson, 1998; Jackson, 1986) and Computational Intelligence (CI) (Bezdek, 1992, 1994; Zurada, Marks, & Robinson, 1994; Pedrycz, 1998) groups. It is believed that the AI group is more appropriate for symbol processing, while the CI group fits more data processing. The CI group contains different methods: neural nets (Haykin, 1994, 2001; Bishop, 1995; Luo & Unbehauen, 1997), statistical pattern recognition (Fukunaga, 1990), fuzzy sets methods (Bezdek, Keller, Krishnapuram, & Pal, 1999), wavelets (Goswami & Chen, 1999), genetic (Mitchell, 1996) and evolutionary algorithms (Bäck, 1996), support vector machines (Cristianini & Shawe-Taylor, 2000), classification and regression trees (Breiman, Friedman, Olsen, & Stone, 1984) and so on.

This chapter considers only *neural nets* methods. Several reasons stand behind the preference given to *neural nets*. These are: 1) one-hidden layer feed-forward *neural nets* have a firm theoretical basis provided by the Kolmogorov's Superposition Theorem (KST) (Kolmogorov, 1957); 2) one-hidden layer Nonlinear Perceptron (NP) can learn a nonlinear mapping more efficiently than any linear network (Barron, 1993, 1994); 3) applied in combination with *clustering* (Duda, Hart, & Stork, 2001; Breiman, 1999), *neural nets* can efficiently learn time-varying, highly multidimensional, nonlinear, and non-stationary data; 4) in spite of common opinion that *neural nets* require utilization of large learning sets and large size of networks (Adeli & Hung, 1995; Adeli & Samant, 2000), there exist several practical ways to significantly mitigate these problems; 5) *neural nets* allow for solving efficiently such important tasks related to modeling (and data mining), as feature selection and visualization.

Consider the advantages of *neural nets* more in detail. There is an old problem of approximation of a continuous real-valued function f of d variables (input dimension) defined on the closed bounded set E (assumed here as a unit hypercube) with a given error ε based on information about P values of the function. For the class of continuous functions the lower bound for P grows exponentially with growth of d . This fact (known under name the "curse of dimensionality") makes reliable approximation of an arbitrary continuous multidimensional function practically impossible for relatively high dimensions. But there exist examples of reliable approximation of functions with high values of d . How may it occur? Modeling a function, one discerns a function from a noise. Actually all methods of modeling explicitly or implicitly assume, that a function has a bounded rate of variability, while the noise may have variability rate higher than that bound. Designers of complex systems often make preliminary statistical system modeling, imitating noise as a statistical distribution with some probability density function (*pdf*). In practice, realizations of a noise with some *pdf* are obtained as continuous functions of a noise with some elementary *pdf*, so-called uniform distribution in the unit interval $[0, 1]$ (Fishman, 1995). The realizations of the uniform distribution are implemented as subsequent values of a continuous piece-wise linear function with very high absolute values of the derivative. Thus, actually a noise is a continuous function with very high rate of variability. In order to discern a function and a noise one has to consider functions from a subclass of the class of continuous functions. Additionally, distribution of a noise is unknown in applications, forcing a designer to choose among several known distributions, such as Gaussian, Weibull (Li & Yu, 1988), and so on, verifying type of distribution using statistical criteria.

Any approximation of a continuous real-valued multidimensional function f can be derived from the Kolmogorov's representation of such a function given by the KST. The KST states that any continuous function of d variables can

21 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/kolmogorovs-spline-complex-network-adaptive/6764

Related Content

Symbolic Function Network: Application to Telecommunication Networks Prediction

George S. Eskander and Amir Atiya (2013). *Artificial Higher Order Neural Networks for Modeling and Simulation* (pp. 237-253).

www.irma-international.org/chapter/symbolic-function-network/71802

Procedure of Medical Diagnosis

(2014). *Medical Diagnosis Using Artificial Neural Networks* (pp. 24-42).

www.irma-international.org/chapter/procedure-of-medical-diagnosis/110997

Impact of Smart Technology and Economic Growth for Biomedical Applications

Dhinesh Kumar R. (2020). *Deep Neural Networks for Multimodal Imaging and Biomedical Applications* (pp. 69-79).

www.irma-international.org/chapter/impact-of-smart-technology-and-economic-growth-for-biomedical-applications/259487

Time Series Data Analysis by Ultra-High Frequency Trigonometric Higher Order Neural Networks

(2021). *Emerging Capabilities and Applications of Artificial Higher Order Neural Networks* (pp. 218-261).

www.irma-international.org/chapter/time-series-data-analysis-by-ultra-high-frequency-trigonometric-higher-order-neural-networks/277679

Qubit Neural Network: Its Performance and Applications

Nobuyuki Matsui, Haruhiko Nishimura and Teijiro Isokawa (2009). *Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters* (pp. 325-351).

www.irma-international.org/chapter/qubit-neural-network/6774