N93-16772

Functional Expansion Representations of Artificial Neural Networks

W. Steven Gray
Assistant Professor
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, Pennsylvania 19104

Abstract

In the past few years significant interest has developed in using artificial neural networks to model and control nonlinear dynamical systems [8]. While there exists many proposed schemes for accomplishing this and a wealth of supporting empirical results, most approaches to date tend to be ad hoc in nature and rely mainly on heuristic justifications. The purpose of this project was to further develop some analytical tools for representing nonlinear discrete-time input-output systems, which when applied to neural networks would give insight on architecture selection, pruning strategies, and learning algorithms. A long term goal is to determine in what sense, if any, a neural network can be used as a universal approximator for nonlinear input-output maps with memory (i.e., realized by a dynamical system). This property is well known for the case of static or memoryless input-output maps [1,5].

The general architecture under consideration in this project was a single-input, single-output recurrent feedforward network described by the n nonlinear difference equations

$$\begin{aligned} x(t+1) &= \sigma(Ax(t)+bu(t)), \qquad x(0)=x_o && (1)\\ y(t) &= cx(t), && (2) \end{aligned}$$

where $\sigma_i \quad i=1,\ldots n$ are the individual activation functions for each neuron, b and c are the input and output connection vectors, A is the recurrent neural network weight matrix, and $x(\cdot)$ is the state vector composed of the outputs of each neuron [4]. For the purpose of modelling input-output behavior we found it convenient to introduce a state-space transformation $z=\sigma^{-1}(x)$ such that the system in (1)-(2) is input-output equivalent to

$$\begin{aligned} z(t+1) &= f(z(t))+g(z(t))u(t), \qquad z(0)=\sigma^{-1}(x_o)\triangleq z_o && (3)\\ y(t) &= h(z(t)), && (4) \end{aligned}$$

where $f(z)=A\sigma(z)$, $g(z)=b$, and $h(z)=c\sigma^{-1}(z)$. The main advantage to using this latter state-space model is that the system is affine in u, and thus, more amenable to analysis by the geometric methods used in nonlinear control theory [2,6]. The basic objective was then to determine a *tractable* functional expansion of the input-output map corresponding to system (3)-(4) in terms of the realization (f,g,h,z_o). A Volterra type expansion

$$y(t)=\sum_{k=0}^{\infty}\sum_{i_1\ldots i_k=1}^{\infty} V_k(i_1,\ldots,i_k)u(t-i_1)\ldots u(t-i_k) \qquad (5)$$

has been published in the literature [4,9], but the expressions for the Volterra kernels, V_k, are too complex for any serious analysis. Thus, we took the alternative approach of deriving a discrete-time analogue of the Fliess functional expansion for general discrete-time nonlinear systems and then applying it in particular to the neural network realization in equations (3)-(4). The resulting expansion was significantly simpler in structure than those published to date and appears promising for future analysis of such systems.

An interesting innovation resulting from this work was the discovery of a natural adaptation of Fliess's work to discrete-time systems. This subject has been studied by other researchers [7,10], but it appears that this particularly tractable approach has been overlooked. One of two main results is described in the following theorem [3]:

Theorem *Suppose the sequence $\{u(k)\}$ is bounded in magnitude on $[0, N]$ and each function of a discrete-time realization (f, g, h, z_o) is analytic on an open subset U of $\Re^n$. Then for sufficiently small N, the input-output mapping can be represented by a convergent generating series*

$$F : u \to y = \sum_{\eta \in I^*} c(\eta) E_\eta, \tag{6}$$

where

$$I^* \text{ is the set of multiindices for the index set } I = \{0, 1\}; \tag{7}$$

$$c(\eta) = c(i_k \ldots i_0) = L_{g_{i_0}} \ldots L_{g_{i_k}} h(z_o) \qquad (g_0(z) \triangleq f(z) - z); \tag{8}$$

$$E_\eta(t) = E_{i_k \ldots i_0}(t) = \sum_{j=0}^{t-1} u_{i_k}(j) E_{i_{k-1} \ldots i_0}(j) \quad t \in [0, N] \tag{9}$$

$$(E_\phi(t) \triangleq 1, \;\; u_0(t) \triangleq 1, \;\; and \;\; u_1(t) \triangleq u(t)). \tag{10}$$

($L_g h$ denotes the Lie derivative of h along g.)

Clearly this result applies to the system (3)-(4). Furthermore, as in the continuous-time case, it is possible to derive a simple series expansion for each Volterra kernel in terms of the coefficients $\{c(\eta) : \eta \in I^*\}$. The details of this analysis are reported in [3].

References

[1] N. E. Cotter, "The Stone-Weierstrass Theorem and Its Application to Neural Networks," *IEEE Trans. on Neural Networks*, vol. 1, no. 4, pp. 290-295, December 1990.

[2] M. Fliess, M. Lamnabhi and F. Lamnabhi-Lagarrigue, "An Algebraic Approach to Nonlinear Functional Expansions," *IEEE Trans. Circuits Syst.*, vol. CAS-30, pp. 554-570, 1983.

[3] W. S. Gray, D. I. Soloway, and O. R. Gonzalez, "Functional Expansions of Nonlinear Discrete-Time Input-Output Mappings," submitted to 1993 IEEE ACC.

[4] N. Z. Hakim, J. J. Kaufman, G. Cerf and H. E. Meadows, "Volterra Characterization of Neural Networks," in Proc. 1991 Asilomar Conference, 1991.

[5] K. Hornik, M. Stinchcombe, and H. White, "Multilayer Feedforward Networks are Universal Approximators," *Neural Networks*, vol. 2, pp. 359-366, 1989.

[6] A. Isidori, *Nonlinear Control Systems*, Berlin: Springer-Verlag, 1989.

[7] S. Monaco and D. Normand-Cyrot, "Nonlinear Systems in Discrete-Time," in *Algebraic and Geometric Methods in Nonlinear Control Theory* (M. Fliess and M. Hazewinkel, eds.), D. Reidel Publishing Company, pp. 411-430, 1986.

[8] K. S. Narendra and K. Parthasarathy, "Identification and Control of Dynamical Systems Using Neural Networks," *IEEE Trans. on Neural Networks*, vol. 1, no. 1, pp. 4-27, 1990.

[9] I. W. Sandberg, "Expansions of Nonlinear Systems," *The Bell System Technical Journal*, vol. 61, no. 2, pp. 159-199, 1982.

[10] E. Sontag, *Polynomial Response Maps*, Berlin: Springer-Verlag, 1979.