The Role of the RKH Space F in the Analysis and Design of Recurrent Neural Networks

Rui J.P. de Figueiredo
Laboratory for Intelligent Signal Processing and Communications
The Henry Samueli School of Engineering
University of California, Irvine
Irvine, CA 92697-2625
Tel: (949)824-7043 Fax: (949)824-2321 e-mail: rui@uci.edu

Abstract - **In this paper, we first briefly recall some of our key results on best approximation in F. These include not only expansions about the origin of H, but also some recent work on expansions about isolated singularities (poles) of f which lead to radial basis functions. We also briefly point out some of the differences between our approach and other similar so-called "kernel-based" approaches, such as the support vector machine (SVM).**
We then move on to general models, based on F, of recurrent neural networks, with and without feedback. In particular, we discuss two specific applications, namely, to nonlinear prediction and to nonlinear digital filter design.

SUMMARY

In this paper we highlight the role that a Reproducing Kernel Hilbert Space (RKHS) F, introduced by us in the late 1970's (see [3]), can play in the modeling, identification, and design of recurrent neural networks.

The space F, and more specifically F(H), is a RKHS of *analytic functionals* on a separable Hilbert Space H (over the field C of complex numbers). F constitutes a weighted Fock Space, a generalization of the conventional symmetric Fock Space, the state space of non-self interacting Boson fields in quantum filed theory.

In our earlier work [1] [2], we introduced F to represent the input-output maps f of large-scale nonlinear dynamical systems, and we showed how such a representation leads to elegant solutions to the problems of optimal modeling, identification, and design of large-scale nonlinear dynamical systems, subject to input-output training data constraints and under the assumption of ellipsoidal models in F for the prior uncertainty in f.

Then in 1990, we showed how the above formulation could be ported to the arena of neural networks (see [3].)

One of the remarkable features of this approach is that a "neural structure" naturally appears in the solution of the underlying non-parametric (infinite-dimensional) optimization problems. In other words, it is not forced a-priori on the formulation of these problems, as, for example, in the case of the error back propagation algorithm.

The second feature is that, on one hand, the representation of f is *very general*. It allows one to describe, exactly or approximately, a very large class of *nonlinear* systems. On the other hand, it enables f, as a member of F, to be processed by *linear* operations, such as the orthogonal projection, pertaining to Hilbert spaces.

Finally our approach enables one to benefit from the conceptual and computational advantages offered by the availability of a Reproducing Kernel for F that can be constructed according to a given application.

In this paper, we first briefly recall some of our key results on best approximation in F. These include not only expansions about the origin of H, but also some recent work on expansions about isolated singularities (poles) of f which lead to radial basis functions. The three different types of expansions are described below.

Since members f of F are analytic functionals on H, they can be expressed as abstract power series (Volterra functional series) in elements x of H converging in an appropriate region Ω in H. Specifically, such a power series can be

Either (a) an expansion about the origin of H (abstract McLaurin series).

$$f(x)=\sum_{n=0}^{\infty}\frac{1}{n!}f_n(x),\quad x\in H \qquad (1)$$

where $f_0(x)=f(0)$ and $f_n(x), n>0$, denotes the Hilbert-Schmidt (H-S) homogeneous polynomial of degree n in elements of H evaluated at x.

Or (b) an expansion about a point x^0 of H (Taylor series):

$$f(x)=\sum_{n=0}^{\infty}\frac{1}{n!}f_n(x-x^0) \qquad (2)$$

where $f_0(x-x^0)=f(x^0)$ and $f_n(x-x^0)$ is defined in the same way as (1) with x replaced by $(x-x^0)$.

Or (c) an expression about a finite number M of points x^i, $i=1,\cdots,M$, considered to be isolated singularities (poles) of f:

$$f(x)=\sum_{i=1}^{M}\ \sum_{n_i=0}^{\infty}\frac{1}{n_i\,!}f_{n_i}(x-x^i) \qquad (3)$$

In each of the above three cases, F can be made a RKHS by endowing it with an appropriate scalar product and equipping it with a Reproducing Kernel that correctly reflects the prior uncertainty in the kernels f_n or more generally $f_{n_i},\ i=1,\cdots,M$.

We then move on to general models, based on F, of recurrent neural networks, with and without feedback. In particular, we discuss two specific applications, namely, to nonlinear prediction and to nonlinear digital filter design.

REFERENCES

[1] R.J.P. de Figueiredo and T.A.W. Dwyer “A best approximation framework and implementation for simulation of large-scale nonlinear systems” (with T.A.W. Dwyer, III), *IEEE Trans. on Circuits and Systems*, vol. CAS-27, no. 11, pp. 1005-1014, November 1980.

[2] R.J.P. de Figueiredo “A generalized Fock space framework for nonlinear system and signal analysis,” *IEEE Trans. on Circuits and Systems*, vol. CAS-30, no. 9, pp. 637-647, September 1983 (Special invited issue on “Nonlinear Circuits and Systems”).

[3] R.J.P. de Figueiredo “Beyond Volterra and Wiener: Optimal Modeling of Nonlinear Dynamical Systems in a Neural Space for Applications in Computational Intelligence: The Experts Speak”, edited by Charles Robinson and David Fogel, volume commemorative of the 2002 World Congress on Computational Intelligence, published by IEEE and John Wiley & Sons, 2003.