Paper:

Feed-Forward Neural Networks Based on the Eigenstates of the Quantum Harmonic Oscillator

Gerasimos Rigatos

Unit of Industrial Automation, Industrial Systems Institute
26504, Rion Patras, Greece
E-mail: grigat@isi.gr

[Received September 1, 2005; accepted January 30, 2006]

The paper introduces feed-forward neural networks where the hidden units employ orthogonal Hermite polynomials for their activation functions. The proposed neural networks have some interesting properties: (i) the basis functions are invariant under the Fourier transform, subject only to a change of scale, (ii) the basis functions are the eigenstates of the quantum harmonic oscillator, and stem from the solution of Schrödinger's diffusion equation. The proposed feed-forward neural networks belong to the general category of nonparametric estimators and can be used for function approximation, system modelling and image processing.

Keywords: feed-forward neural networks, quantum harmonic oscillator, Schrödinger's diffusion equation, Gauss-Hermite expansion, Hermite polynomials

1. Introduction

Feed-forward neural networks (FNN) are the most popular neural architectures due to their structural flexibility, good representational capabilities, and availability of a large number of training algorithms. The hidden units in a FNN usually have the same activation functions and are usually selected as sigmoidal functions or gaussians. This paper presents feed-forward neural networks that use orthogonal Hermite polynomials as basis functions. The proposed neural networks have some interesting properties: (i) the basis functions are invariant under the Fourier transform, subject only to a change of scale (ii) the basis functions are the eigenstates of the quantum harmonic oscillator (QHO), and stem from the solution of Schrödinger's diffusion equation. The proposed neural networks belong to the general category of nonparametric estimators and are suitable for function approximation, system modelling and image processing. Two dimensional feed-forward quantum neural networks can be also constructed by taking products of the one-dimensional basis functions.

Feed-forward neural networks that use the eigenstates of the quantum harmonic oscillator (Hermite basis functions) demonstrate the particle-wave nature of information as described by Schrödinger's diffusion equation [1, 2]. Attempts to enhance the connectionist neural models with quantum mechanics properties can be also found in [3–6]. The proposed FNNs extend previous results on neural structures compatible with quantum mechanics postulates, given in [7, 8].

In this paper, it is considered that the input variable x of the neural network can be described not only by crisp values (particle equivalent) but also by the normal modes of a wave function (wave equivalent). Since the basis functions of the proposed FNN are the eigenstates of the quantum harmonic oscillator, the FNN's output will be the weighted sum $\psi(x)=\sum_{i=1}^{n} w_k \psi_k(x)$, where $|\psi(x)|^2$ is the probability that the input of the neural network (quantum particle equivalent) is found between x and $x+\Delta x$. Thus, the weight w_k provides a measure of the probability to find the input on the neural network in the region associated with the eigenfunction $\psi_k(x)$.

The structure of the paper is as follows: In Section 2 an overview of feed-forward neural networks is presented. In Section 3, the solution and the eigenstates of the quantum harmonic oscillator are summarized. In Section 4 the expansion in Gauss-Hermite series is explained and the feed-forward neural networks which are based on the eigenstates of the quantum harmonic oscillator are introduced. In Section 5 simulation results on the capabilities of the proposed FNNs for function approximation, system modelling and image processing are given. Finally, in Section 6 concluding remarks are stated.

2. Feed-Forward Neural Networks

Feed-forward neural networks (FNN) serve as powerful computational tools, in a diversity of applications including function approximation, classification and pattern recognition. When equipped with procedures for learning from measurement data they can generate models of unknown systems. Feed-forward neural networks are the most popular neural architectures due to their structural flexibility, good representational capabilities, and availability of a large number of training algorithms.

The idea of function approximation with the use of feed-forward neural networks (FNN) comes from generalized Fourier series. It is known that any function $\psi(x)$

in a L^2 space can be expanded in a generalized Fourier series in a given orthonormal basis, i.e.

$$\psi(x)=\sum_{k=1}^{\infty}c_k\psi_k(x),\ a\le x\le b. \qquad \ldots\ldots \tag{1}$$

Truncation of the series yields in the sum

$$S_M(x)=\sum_{k=1}^{M}a_k\psi_k(x). \qquad \ldots\ldots\ldots\ldots \tag{2}$$

If the coefficients a_k are taken to be equal to the generalized Fourier coefficients, i.e. when $a_k=c_k=\int_a^b\psi(x)\psi_k(x)dx$, then Eq.(2) is a mean square optimal approximation of $\psi(x)$.

Unlike generalized Fourier series, in FNN the basis functions are not necessarily orthogonal. The hidden units in a FNN usually have the same activation functions and are often selected as sigmoidal functions or gaussians. A typical feed-forward neural network consists of n inputs $x_i,\ i=1,2,\ldots,n$, a hidden layer of m neurons with activation function $h:\ R\rightarrow R$ and a single output unit (see **Fig.1**). The FNN's output is given by

$$\psi(x)=\sum_{j=1}^{n}c_jh\Big(\sum_{i=1}^{n}w_{ji}x_i+b_j\Big). \qquad \ldots\ldots \tag{3}$$

The root mean square error in the approximation of function $\psi(x)$ by the FNN is given by

$$E_{RMS}=\sqrt{\frac{1}{N}\sum_{k=1}^{N}(\psi(x^k)-\hat{\psi}(x^k))^2} \qquad \ldots\ldots \tag{4}$$

where $x^k=[x_1^k,x_2^k,\ldots,x_n^k]$ is the k-th input vector of the neural network. The activation function is usually a sigmoidal function $h(x)=\frac{1}{1+e^{-x}}$ while in the case of radial basis functions networks it is a Gaussian. Several learning algorithms for neural networks have been studied. The objective of all these algorithms is to find numerical values for the network's weights so as to minimize the mean square error E_{RMS} of Eq.(4). The algorithms are usually based on first and second order gradient techniques. These algorithms belong to: i) batch-mode learning, where to perform parameters update the outputs of a large training set are accumulated and the mean square error is calculated (back-propagation algorithm, Gauss-Newton method, Levenberg-Marquardt method, etc.), ii) pattern-mode learning, in which training examples are run in cycles and the parameters update is carried out each time a new datum appears (Extended Kalman Filter algorithm).

3. Eigenstates of the Quantum Harmonic Oscillator

Feed-forward neural networks with Hermite basis functions show the particle-wave nature of information, as described by Schrödinger's diffusion equation, i.e.

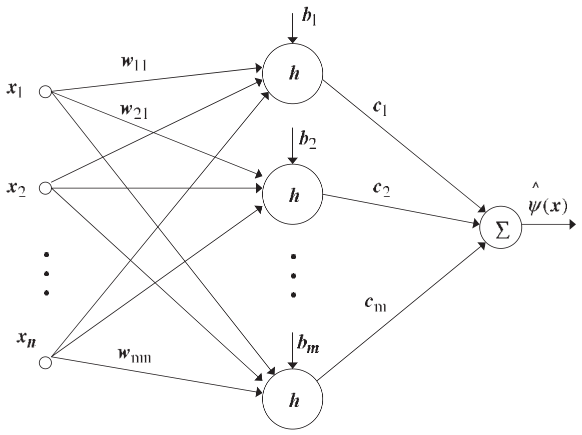
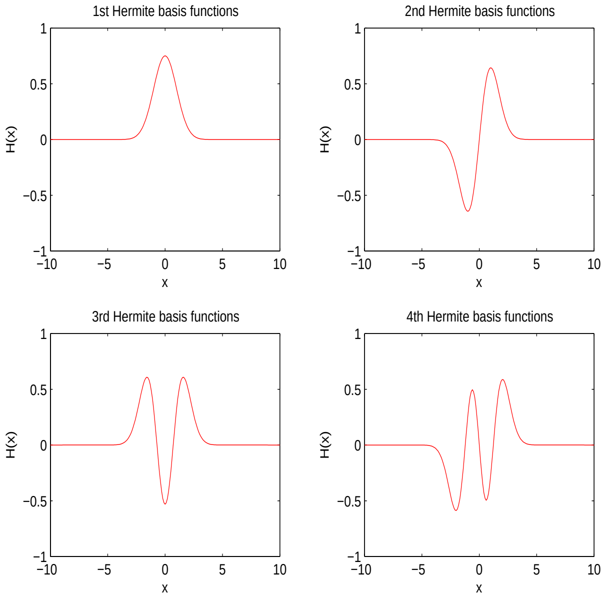


Fig. 1. Feed-forward Neural Network.


1st Hermite basis functions

2nd Hermite basis functions

3rd Hermite basis functions

4th Hermite basis functions

Fig. 2. Analytic diagrams of one-dimensional Hermite basis functions.

$$i\hbar\frac{\partial\psi(x,t)}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\psi(x,t)+V(x)\psi(x,t)$$
$$\Rightarrow i\hbar\frac{\partial\psi(x,t)}{\partial t}=H\psi(x,t) \tag{5}$$

where $\hbar$ is Planck's constant, H is the Hamiltonian, i.e. the sum of the potential $V(x)$ and of the Laplacian $-\frac{\hbar^2}{2m}\nabla^2=-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$. The probability density function $|\psi(x,t)|^2$ gives the probability at time instant t the input x of the neural network (quantum particle equivalent) to have a value between x and $x+\Delta x$. The general solution of the quantum harmonic oscillator, i.e. of Eq.(5) with $V(x)$ being a parabolic potential, is [1, 2]:

$$\psi_k(x,t)=H_k(x)e^{-x^2/2}e^{-i(2k+1)t}\quad k=0,1,2,\ldots \tag{6}$$