ESANN'2008 proceedings, European Symposium on
Artificial Neural Networks - Advances in Computational Intelligence and Learning
Bruges (Belgium), 23-25 April 2008, d-side publi., ISBN 2-930307-08-0.

A Novel Autoassociative Memory on the Complex Hypercubic Lattice

Rama Murthy Garimella and Praveen Dasigi

International Institute of Information Technology (IIIT-H),
Hyderabad - India

Abstract. In this paper we have defined a novel activation function called the multi-level signum for the real and complex valued associative memories. The major motivation of such a function is to increase the number of patterns that can be stored in a memory without increasing the number of neurons. The state of such a network can be described as one of the points that lie on a complex bounded lattice. The convergence behavior of such a network is observed which is supported with the simulation results performed on a sample dataset of 1000 instances

1 Introduction

The autoassociative memory first pioneered by Hopfield, has proved that a fully connected network with the neurons implementing the SIGNUM activation function is capable of recalling a pattern that is stored previously in the network given a spurious or a noisy pattern [1]. The associated convergence theorem was proven stating that the network converges to a stable state when it operates in the serial mode and to a cycle of length ≤ 2 when it operates in the parallel mode [3]. The neuron in the associative memory computes the following activation function. If $\{x_i, 1 \leq i \leq M\}$ are inputs to a neuron and $\{w_i, 1 \leq i \leq M\}$ are the synaptic weights, then the output of the neuron is given by

$$y = sign \sum_{j=1}^{M} [w_j x_j - T],$$

where,

$$sign(z) = \begin{cases} 1, & z \geq 0 \\ -1, & z < 0 \end{cases}$$

Here, T is the threshold at a neuron. It is easily seen that there is no loss of generality in assuming T is equal to zero.

In the context of design of Complex Valued Neural Networks (CVNNs), Aizenberg, Zurada et al., proposed a novel activation function using the concept of "Phase Quantization" [4, 5] (of the complex valued net contribution at a neuron). The complex-valued neural network is an associative memory where the inputs to the neurons are complex numbers and the synaptic weights pertaining to the network are complex as well [7]. As the network operates, multiplication of general complex-valued weight w mixes the real and imaginary quantities in a certain manner determined by the value w. In the network, the multiplication of w is executed for all the parallel input data elements x_i. As a result, the output maintains a certain vector-direction relation in the complex plane. This is one of the properties that we can utilize effectively to treat two-dimensional information In an effort to innovate complex valued associative memory, the authors proposed a novel complex signum function in [6]. The phase quantization is bounded but has the flexibility of increasing the number of states in the quantization. To perform the quantization in Argand plane, the constraint is that we apply bounds on the quantization. This has led the authors to explore the idea of a bounded spatial quantization scheme to increase the number of states in the real as well as the complex valued networks.

The paper is organized as follows. Section 2 will give a brief introduction to the general functioning of the activation functions on the hypercubic lattice, which is the major contribution of this paper. Section 3 will discuss the convergence issues regarding the neural nets on the hypercubic lattice in the real and complex domains and putforth some supporting simulation results along with potential applications. Section 4 will conclude the work

2 Activation Functions

The signum activation function as explained in the first section will change the state of the neuron to ± 1 based on the value of the weighted sum of inputs. In an associative memory, with individual units being such neurons, the state of the network can be depicted as a vector $X = (x_1, x_2, ..., x_n)$ where $x_i = \pm 1$. Each vector X can be visualized as the vertex of a hypercube of order n. Thus the network state moves through the vertices of the hypercube till it reaches a state where the energy of the network is minimum. In this setup, the total number of states that the network can possibly stay at will be 2^n. The classical Hopfield network [1] that uses this setup is an auto associative memory that retrieves stored patterns, given spurious or noisy patterns as input. The number of patterns that can be stored is limited due to the fact that the possible values that each element of the pattern can take will be ± 1. In this paper we have experimented with the concept of increasing the number of patterns that can be stored by proposing a multilevel signum function. The multi-level signum function $mlsign(x)$, on giving an input x, quantizes the weighted sum to one of L predefined states between $[-1, 1]$. To do so would be not as straightforward as the operation of the real valued signum, since the output has to be bounded both ways and within the range $[-1, 1]$. Thus, before operating the multi-level function we need to normalize the weighted sum so that it lies between the range $[-1, 1]$. The output of the normalization step will be further fixed to the nearest state of the L predefined states. The requirement of the normalization step is to bring the weighted sum to a fixed range consistently. One way to do this is to divide the output with the difference between the maximum and minimum of the range of weighted sums at that instant. This would ensure that the output is between $[-1, 1]$. However an even consistent way to do this is to use the hyperbolic tangent function to bring the weighted sum within the required range. After this step the quantization can be done by fixing the output

ESANN'2008 proceedings, European Symposium on
Artificial Neural Networks - Advances in Computational Intelligence and Learning
Bruges (Belgium), 23-25 April 2008, d-side publi., ISBN 2-930307-08-0.

values as suggested. Thus, for the specified number of levels L for an input z, the function will be,

$$y = \begin{cases} l_1, & l_1 \leq tanh(z) < l_2 \\ l_2, & l_2 \leq tanh(z) < l_3 \\ \vdots & \\ l_{L-1}, & l_{L-1} \leq tanh(z) < l_L \end{cases}$$

In this specific case however, we have $l_1 = -1$ and $l_L = 1$. For n neurons the use of this activation function thus increases the number of states in the network from 2^n to L^n. The network will thus lie on a structure called the hypercubic lattice. The hypercubic lattice is a set of points within the hypercube $X = \{x_1, x_2, ..., x_n\}$ where x_i can take L uniformly spaced integer values between $[-1, 1]$. Thus the network proceeds along the points in the hypercubic lattice till it finds the stable state with the maximum energy. The setup described above is the basis for the complex valued multilevel signum activation function. The complex valued autoassociative memory is defined in a previous work by the authors and is proved to converge to a stable state in [6]. In the complex valued autoassociative memory the input is a vector of complex numbers, each of which is from the set$\{1 + \jmath, 1 - \jmath, -1 + \jmath, -1 - \jmath\}$ The weights are complex and can be constituted in a hermitian matrix. The activation function defined at the neuron is the complex-valued signum as follows,

$$CSGN(z) = [sign(\Re(z))] + \jmath[sign(\Im(z))]$$

Here, $\Re(z)$ and $\Im(z)$ describe the real and imaginary parts of the complex number z. The output of each of the neuron leads to a vector X, the current state of the network which lies on what is defined to be the complex hypercube. This formulation can be extended to increase the number of states in the complex-valued associative memory, leading to what can be termed as a complex hypercubic lattice.

$$CMLSIGN(z) = [fix(tanh(\Re(z)))] + \jmath[fix(tanh(\Im(z)))]$$

Here the function $fix()$ rounds the input to the nearest level less than or equal to the input as described in 1. The complex valued auto associative memory with complex bi-level signum moves over the vertices of the complex hypercube till it finds the state with the maximum energy value. At each state it picks the neighboring vertex that contributes to a higher energy value. In a similar way, the associative memory that uses the multilevel signum in the real and complex case moves over the points in the hypercubic lattice in both the real and the complex cases. The convergence properties such memories will be discussed in detail in the next section

3 Convergence behavior on the hypercubic lattice

The content addressable memory first designed by Hopfield [1] is based on the formulation that the activation function leads to an overall energy which follows a monotone decrease to a stable least energy state. A generalized proof has been offered by Bruck et al [3]., where the formulation is that the energy is non- decreasing and reaches a global maximum eventually. In every standard case of the associative memory, the energy monotonically reaches a stable state by "sliding" over the vertices of the hypercube. It has been proven that similar behavior is exhibited by a complex valued associative memory on the complex hypercube in [6]. The associative memory can operate in one of the two modes,i.e., serial or parallel. In the serial mode the activation function is computed at one neuron at each step and the energy is updated. In the parallelmode, the activation function takes into account all the neurons simultaneously and the energy is updated. The classic autoassociative memory in the real [3] and the complex [6] case will converge to a single stable state in the serial mode and to a cycle of length ≤ 2 in the fully parallel mode. If we take into consideration the the difference in energy in the classic autoassociative case in the serial mode, currently at a neuron k, it can be described by the equation

$$\Delta E = E(t+1) - E(t) = 2\Delta V_k H_k + W_{k,k} {\Delta V_k}^2$$

where,

$$H_i(t) = \sum_{X_j=1}^{n} W_{j,i} V_j(t) - T_i$$

In the equation for ΔE, the expression V_k which is the difference in the input from $t+1$ to t, and H_k will be of same sign according to the definition of the activation function. Hence ΔE is always positive. The same holds true for the complex valued associative memory as proved in [6], using a novel proof technique by considering the real and imaginary component separately and proving them to converge. However when it comes to the case of the multi-level signum function, this case will not hold.

$$\textbf{if } V_k(t+1) = l_p, \Delta V_k . H_k \begin{cases} \geq 0, & \textbf{if } V_k(t) \in [l_1, l_p] \textbf{ and } l_p \geq 0 \\ & \textbf{if } V_k(t) \in [l_p, l_L] \textbf{ and } l_p < 0 \\ < 0, & \textbf{if } V_k(t) \in (l_p, l_L] \textbf{ and } l_p \geq 0 \\ & \textbf{if } V_k(t) \in [l_1, l_p) \textbf{ and } l_p < 0 \end{cases}$$

Thus in the case of an associative memory on the hypercubic lattice, the energy function cannot be a monotonic one. This is due to the reason that as it proceeds along the points of the lattice, at each point it will have to chose between neighbors that might have a similar energy due to the increase of states. To observe this from a practical perspective, some simulation results are to be looked at.

3.1 Simulation results

Some simulations have been performed to test the convergence of energy for the real and complex valued associative memories on the hypercubic lattice. For the

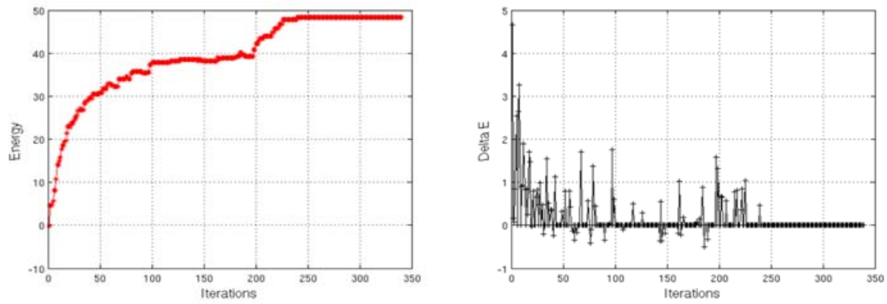

Fig. 1: (a) shows the convergence of the energy function to a global maximum in the complex valued case, (b) shows the convergence of ΔE to zero as the energy reaches a global maximum

simulations we have generated 1000 random weight matrices for the real and complex case, at different sizes ranging from 10 to 150 neurons in each network. It has been observed from the simulations that when operating in the serial mode, the network always converges to a stable state and for the parallel mode converges to a cycle of length ≤ 2, which is similar in behavior to the classic hopfield network. The explanation for this behavior is that when the network operates in the serial mode, each neuron can take a total of L values, and thus the value at one of the nodes can affect the energy to have a spurious temporary divergence from the monotone increase. However it still increases after a certain state to achieve convergence to a global maximum. This can be observed in the figures shown below.

As we can see in the figure 1 the energy sometimes seems to enter a spurious energy state which is less than the previous state, however it leaves the state again to reach a global maximum. This behavior is repeated for every configuration tested. The authors are currently in the process of discovering a proof technique to prove the global convergence behavior of such associative memories. It is expected that the proof arguments and techniques that are utilized in non-linear dynamics [8] will be applicable to the neural network that is being considered here.

3.2 Applications

The major motivation behind the design of this novel associative memory is to explore the capability of a network that can store and recall a greater variety of patterns. This can be made possible since the the number of states for each neuron increases. To test the capability of the network, one needs to also check the stability with higher variety of patterns. Increase in the number of states to a certain extent will increase the capacity of the network, though it may lead to ambiguous and spurious outputs after a certain increase. In any case, a network with increased capability of storage can be employed to yield better results in image and video restoration and coding. The plausibility of such applications is the target for future research.

4 Conclusions and Future work

In this paper we have defined a neural network configuration that operates on a novel multi-dimensional structure called the complex hypercubic lattice. The convergence behavior of such a network has been studied which is supported by the simulation results putforth. The motivation is to increase the number of patterns that can be stored in an auto associative memory in the real and complex-valued cases. The immediate future work is to analyze the behavior of the network and bring out a formal proof to the convergence characteristics. Further the storage and recall capacity of such a network has to be analyzed.

References

[1] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci.* USA 79: 2554-2558 (1982).

[2] J.J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons.*Biophysics*, 79:2554-2558, April 1982

[3] Jehoshua Bruck, Joseph W. Goodman, A generalized convergence theorem for neural networks *IEEE Transactions on Information Theory*, 4(5): 1089- (1988)

[4] S. Jankowski, A. Lozowski, and J. M. Zurada, Complex-valued multistate neural associative memory, *IEEE Trans. Neural Networks*, vol. 7, no. 6, pp. 1491-1496, November 1996.

[5] M. K. Muezzinoglu and J. M. Zurada, A new design method for complex-valued multistate Hopfield associative memory, *proceedings of the International Joint Conference on Neural Networks*, Portland, OR, July 20-24, 2003, vol. 1, pp. 45-50.

[6] G.Ramamurthy and D.Praveen, Complex-valued Neural Associative Memories on the Complex Hypercube, *proceedings of the IEEE conference on Cybernetics and Information Systems* (CIS 2004), Singapore, Dec 2004.

[7] A. Hirose. *Complex valued Neural Networks: Theories and Applications*, World scientific publishing Co, November 2003.

[8] Steven H Strogatz. *Non-linear Dynamics and Chaos*, Perseus Publishing. ISBN 0-7382-0453-6.