濟大論文集(自然), 37, 301-306, 1993
Cheju Univ. Jour. (Natural Sci.),
37. 301-306. 1993

Space Analysis of Hopfield Neural Network

*Min-Je Kang**

호프필드 신경회로망의 공간분석

강 민 제*

Summary

The Hopfield network has been used in solving the optimization problems required of much calculations. Because the Hopfield network computes rapidly this kind of problem using its parallel processing property. In this paper, the energy function has been analyzed to help understanding the properties of this network. It has been proved here that the energy function of the Hopfield network has one stationary point which is a saddle point in the unconstrained space, therefore, this energy function has the constrained minima on the boundary of the constrained space.

Introduction

Artificial neural systems are densely interconnected systems which have been developed with the hope of simulating performance that can eventually match that of the human brain. They are inspired by biological systems which perform the task of pattern recognition much more quickly than computers by using components much slower than those found in computers. In 1982, Hopfield proposed a simplified electrical model of the fully connected neural network and showed the network's computational properties. Since then, many researchers have begun further study to exploit the algorithms and applications of the Hopfield model. The network computes a solution by following a path that decreases the Liapunov-like energy, just as a rain drop moves downhill to minimize its gravitational potential energy. If the energy function of the Hopfield network contains more than one minimum, the network will converge to the minimum nearest the initial point. This means that the network converges to a local minimum rather than the global minimum.

A new mathematical analysis has been developed for the energy function of the Hopfield network. This method provides a detailed insight

* 공과대학 전자공학과(Dept. of Electronics Engineering, Cheju Univ., Cheju-do, 690-756, Korea)

into the transient behavior and stability conditions of the network.

The Neural Network Formulation

The Hopfield network is a simple neural circuit which consists of synaptic connections and neurons. In the actual network, the non-linear input-output relation of neuron is determined by a sigmoid function

$$V_i = \frac{1}{1+e^{-\lambda u_i}} \quad (1)$$

The neurons are coupled together by a set of non-linear differential equations,

$$C_i \frac{du_i}{dt} = \sum_{j=1}^{n} W_{ij}V_j + I_i - u_iG_i, \quad \text{for } i=1, \cdots, n \quad (2)$$

where n is the number of neurons, V_i is the neural output, u_i is the neural input, w_{ij} is the synaptic connection, G_i is the total conductance to input node, and C_i is the input capacitance. Hopfield showed that provided $W_{ij}=W_{ji}$ and $W_{ii}=0$ for all i and j, the state of the network v_i converges to a local minimum of a Lyapunov energy function

$$E(v) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}V_iV_j - \sum_{i=1}^{n} I_iV_i + \sum_{i=1}^{n} G_i \int_{0.5}^{v_i} f_i^{-1}(z)\,dz \quad (3)$$

It is claimed that $dE/dt<0$ for all i, therefore, it shows that the Lyapunov energy function will always be minimized (Hopfield, 1984).

Existence of Minima of the Energy Function with the Infinite Sigmoid Gain

Proposition : *If the sigmoid gain λ is infinite, the connection conductance matrix $\underline{W}$ is symmetric, and the diagonal elements are zeros, (i. e. $w_{ii}=0$), then the following statement is true :*

The energy function always has a saddle point in the unconstrained space.

Proof : Finding the equilibrium point of the postulated energy function $E(\underline{v})$ in equation (3) takes the following form

or, briefly

$$\underline{\nabla}E(\underline{v}) = 0 \quad (4)$$

$$-\underline{W}\,\underline{v} - \underline{i} + \underline{G}\,\underline{u} = 0 \quad (5)$$

where

$$u = \frac{1}{\lambda}\ln\left(\frac{v}{1-v}\right) \quad (6)$$

Since the sigmoid gain λ is infinite, the last term of equation (5) will disappear, so

$$-\underline{W}\,\underline{v} - \underline{i} = 0 \quad (7)$$

It is well-known that if matrix $\underline{A}$ is symmetric, then $\underline{A}$ is nonsingular so that its inverse matrix exists (Friedberg, 1986). Therefore, $\underline{W}$ is nonsingular, since $\underline{W}$ is designed to be symmetric for this type of network. Because of nonsingularity of $\underline{W}$, only the single stationary point exists and is given as follows

$$\underline{v}^* = -\underline{W}^{-1}\,\underline{i} \quad (8)$$

The Hessian matrix of $E(\underline{v})$ can be used to determine if the stationary point is a minimum, a maximum, or a saddle point. Its value is

$$\underline{\nabla}^2 E(\underline{v}) = -\underline{W} \quad (9)$$

Since the diagonal elements of the Hessian matrix (9) are all zero (because $w_{ii}=0$ for all i) and the Hessian matrix is symmetric, the Hessian matrix is always indefinite. Thus, the stationary point obtained from equation (8) must be a saddle point. This will be explained below.

Since the sum of all eigenvalues of a matrix equals its trace (Friedberg, 1986) and the trace of the Hessian matrix is zero ($\sum_{i=1}^{n} w_{ii}=0$) for the Hopfield networks, the sum of all eigenvalues of the Hessian matrix is zero. This means that all eigenvalues are zero, or that some eigenvalues are negative and the others are positive for the summation to be zero. Besides, all eigenvalues of the Hessian matrix cannot be zero and this will be explained.

It is a known result that if $\underline{A}$ is a symmetric matrix, then a diagonal matrix can be found which is similar to $\underline{A}$, and similar matrices have the same traces (Friedberg, 1986) and eigenvalues (Friedberg, 1986). Since the Hessian matrix is symmetric (because the connection matrix $\underline{W}$ is designed to be symmetric), the Hessian matrix is similar to a diagonal matrix. This diagonal matrix does not have zero elements on its diagonal because similar matrices have the same rank. Furthermore, all elements of this diagonal matrix are eigenvalues of both matrices (Ortega, 1972). Thus, the Hessian matrix does not have zero eigenvalues. Therefore, since the summation of all eigenvalues yields zero, it is obvious that some eigenvalues are negative and the others are positive which proves that the Hessian matrix is indefinite.

This proves completely that the energy function always has a distinct stationary point and that this stationary point is a saddle point. The following case study illustrates that the energy function always has a saddle point in the unconstrained space.

Case Study

The 3-bit A/D converter (Tank and Hopfield, 1986) is selected as a case study to evaluate the existence of stationary points. This case study is used to verify that the ideal energy function always has a stationary point and this point is a saddle point.

The energy function with an infinite sigmoid gain λ can be expressed by substituting the weights and the bias current for the 3-bit A/D converter as follows

$$E=-\frac{1}{2}\,[v_1\ v_2\ v_3]\begin{bmatrix}0 & -2 & -4\\ -2 & 0 & -8\\ -4 & -8 & 0\end{bmatrix}\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix} -[v_1\ v_2\ v_3]\begin{bmatrix}x-\frac{1}{2}\\ 2x-2\\ 4x-8\end{bmatrix} \quad (10)$$

To find a stationary point of $E(\underline{v})$, the gradient of the energy function can be used

$$\underline{\nabla}E(\underline{v})=-\begin{bmatrix}0 & -2 & -4\\ -2 & 0 & -8\\ -4 & -8 & 0\end{bmatrix}\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}-\begin{bmatrix}x-\frac{1}{2}\\ 2x-2\\ 4x-8\end{bmatrix}=\underline{O} \quad (11)$$

It is also instructive to see that the Hessian matrix $\underline{H}$ of the energy function becomes

$$\underline{\nabla}^2E(\underline{v})=\begin{bmatrix}0 & 2 & 4\\ 2 & 0 & 8\\ 4 & 8 & 0\end{bmatrix} \quad (12)$$

The eigenvalues of $\underline{H}$ in (12) are obtained as -8.2788, -1.5699 and 9.8487. Since the signs of all eigenvalues differ (two eigenvalues are negative and one is positive), $\underline{H}$ is indefinite. Furthermore, sum of these eigenvalues is zero as we expected. This is, as we explained before, because $\underline{H}$ is symmetric and all diagonal elements of $\underline{H}$ are zero.

Thus, whatever solution $\underline{v}^\star$ of (8) is identified such that

$$\underline{\nabla}E(\underline{v}^\star)=0 \quad (13)$$

it will be a saddle point of the energy function.

The solution for a saddle point of the energy function can be obtained by using equation (8). Since $\underline{W}$ is symmetric, its inverse matrix exists as follows

$$\underline{W}^{-1} = -\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{8} \\ -\frac{1}{4} & \frac{1}{8} & -\frac{1}{16} \\ -\frac{1}{8} & -\frac{1}{16} & \frac{1}{32} \end{bmatrix} \tag{14}$$

Therefore, the distinct saddle point exists as follows

$$\begin{bmatrix} v_1^* \\ v_2^* \\ v_3^* \end{bmatrix} = -\begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & -\frac{1}{8} \\ -\frac{1}{4} & \frac{1}{8} & -\frac{1}{16} \\ -\frac{1}{8} & -\frac{1}{16} & \frac{1}{32} \end{bmatrix} \begin{bmatrix} x-\frac{1}{2} \\ 2x-2 \\ 4x-8 \end{bmatrix} \tag{15}$$

Expanding equation (15) results in

$$\begin{bmatrix} V_1^* \\ V_2^* \\ V_3^* \end{bmatrix} = \begin{bmatrix} \frac{x}{2} - \frac{5}{4} \\ \frac{x}{4} - \frac{3}{8} \\ \frac{x}{8} + \frac{1}{16} \end{bmatrix} \tag{16}$$

The saddle point can be obtained for selected x because equation (16) is a function of analog input x in this case. To obtain the locus of the saddle points, two independent equations can be obtained by removing x from equation (16) as follows

$$\begin{aligned} v_2 &= \frac{1}{2} v_1 + \frac{1}{4} \\ v_3 &= \frac{1}{2} v_2 + \frac{1}{4} \end{aligned} \tag{17}$$

Since each equation has only two variables in a three dimensional space, these are plane equations. Therefore, the intersection of these two planes is the locus of the saddle point. The surface in Fig. 1. is drawn from the first equation in (17), and the line on the surface is the intersection of the above two equations.

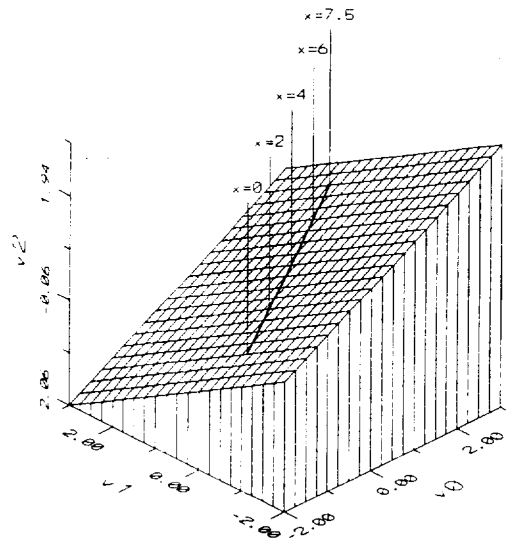

Fig. 1. Loocus of Saddle Points for the 3-bit A/D Converter

Conclusion

In this paper, we have given a complete analysis of the energy function of Hopfield network with infinite sigmoid gain. The energy function of Hopfield network with two state neurons has only one stationary point which is saddle point in unconstrained space. Therefore, the hopfield network has the constrained minimum on the boundary of the constrained space. This paper has proven that this proposition agrees with many experimental results that the Hopfiled network with two state neurons converges to the boundary of the hypercube.

References

Hopfield, J. J., 1984. Neurons with graded response have collective computational properties like those of two state neurons, *Proc. Natl. Acad. Sci.* 81 : 3088-3092.

Ortega, J. M., *Numerical Analysis*, Acdemic Press, New York, 1972.

Friedberg and Stephen, *Linear Algebra with Applications*, Prentice-Hall, Englewood Cliffs, N. J., 1986.

Tank, D. W. and J. J. Hopfield, 1986. Simple neural optimization networks : An A/D converter, signal decision circuit and a linear programming circuit, *IEEE Trans. on Circ. and Syst.* CAS-33(5) : 533-541.

Zurada, J. M. and M. J., Kang, 1989. Computational circuits using neural optimization concept, *Int. Journal of Electron.*, 67(3) : 311-320.

〈국문초록〉

호프필드 신경회로망의 공간분석

호프필드 신경회로망의 에너지 함수를 분석하여 신경회로망의 안정상태와 수렴과정을 쉽게 이해할 수 있다. 호프필드 신경회로망의 에너지 함수는 전 공간에서 오직 하나의 변곡점만을 갖으므로 즉, 다른 극점이 존재하지 않으므로, 제한된 공간에서만 제한된 극소점이 존재하게 되어서 신경회로망은 그 점으로 수렴하게 된다. 이 논문에서는 호프필드 신경회로망의 에너지 함수는 전 공간에서 다른 극점이 없고 오직 변곡점만이 존재함을 증명하였다.