Convergence analysis of recurrent neural network with self-loops based on eigenvalues of a connection matrix

Convergence Analysis of Recurrent Neural Network with Self-loops Based on Eigenvalues of A Connection Matrix

Yoshihiro Tomikawa
Graduate School of Natural Sci.& Tech., Kanazawa Univ.
2-40-20,Kodatsuno, Kanazawa, 920 Japan
E-mail tomy@rd.ykk.co.jp

Kenji Nakayama
Faculty of Eng., Kanazawa Univ.
2-40-20,Kodatsuno, Kanazawa, 920 Japan
E-mail nakayama@ec.Kanazawa-u.ac.jp

ABSTRACT

Recurrent Neural Network(RNN) can be applied to solve a combinatorial optimization problem. However, the existence of the local minima in this model prevents its application to real world. In this paper, an analysis method for network dynamics based on eigenvalues and eigenvectors of a connection weight matrix is proposed. In this analysis, the transition of the number of negative eigenvalues and the movement of eigenspaces by increasing diagonal elements,which correspond to feedback loops of RNN, are discussed. From this analysis, it is confirmed that the number of negative eigenvalues decreases and the eigenspaces move toward the ascent side of the energy slope at the center point of the state space of RNN,as increasing diagonal elements. These behaviors of RNN contribute to the improvement for searching a solution of a combinatorial optimization problem.

This analysis method is applied to a 2-dimensional example and 5 cities T.S. problems. From this analysis, it is theoretically found that the network performance of detecting the optimal solution can be improved by increasing the values of diagonal elements of connection matrix.

1. Introduction

A combinatorial optimization problem is a class of NP-complete problems. Hopfield and Tank[1] suggested that a Recurrent Neural Network(RNN), which does not use self-feedback loops, can provide plausible solutions for such a problem with less number of searches. Since this kind of problem can be applied to a variety of applications, a lot of researches have been done[1]-[8]. However, since the above RNN has local minima, its solution is not always guaranteed to be the global optimal one. This defect prevents the use of the RNN to real applications.

For the improvement of the convergence probability to the optimal solution, Uesaka [6] proposed the method of controlling the initial state of RNN. However, this method is still experimental and is theoretically proved.

Besides, Abe et al[5]and Ohta et al[7] proposed the nonzero diagonal element method. The diagonal elements of a connection weight matrix correspond to self-feedback loops of neurons in a RNN. Then, the nonzero diagonal element means that

RNN uses self-loops. They suggested that the network performance for a TSP problem could be improved by uniformly increasing the diagonal elements. However, this improvement has not been well analyzed theoretically.

On the other hand, Aiyer et al[8] proposed that the eigenvalue analysis can be applied to RNN dynamics. Since the eigenvalues of a connection weight matrix are related to the network behavior, they provide available information about the network dynamics.

In this paper, we extend the eigenvalue analysis method to convergence analysis of a RNN with feedback loops. That is, the effects of diagonal elements on the network dynamics are considered.

We employ an analog model for this analysis. Then the convergence states are not guaranteed to take binary values. This is different from a discrete model like the Hopfield model. However, both models are same in the viewpoint that they can change their states in the direction of decreasing the energy function. Then, discussions in this paper may be valid for the discrete model.

2. Energy Function Based on Eigenvalues and Eigenvectors

2.1. Energy Function of RNN with Feedback Loops

The energy function $E(\mathbf{x})$ of a RNN is given as the following Nth-order quadratic form.

$$E(\mathbf{x}) = \frac{1}{2}\mathbf{x}^t\mathbf{W}\mathbf{x} + \mathbf{th}^t\mathbf{x} \tag{1}$$

where $\mathbf{x}$ is the state vector of the network, $\mathbf{W}$ is a connection weight matrix and $\mathbf{th}$ is a threshold vector. We consider the case that the diagonal elements of the matrix $\mathbf{W}$ can take non-zero values. The diagonal elements mean the self feedback loops of RNN. In order to express the change of diagonal elements explicitly, the connection weight matrix and the threshold vector can be written as the following equations.

$$\mathbf{W} = \mathbf{W}_0 + 2a\mathbf{I} \tag{2}$$

$$\mathbf{th} = \mathbf{th}_0 - a\mathbf{i} \tag{3}$$

where $\mathbf{I}$ is a unit matrix and $\mathbf{i}$ is a vector which has only components of 1. In these equations, a means the uniform change of diagonal elements. The letters with the suffix 0 mean the invariant matrix and vector by the change of diagonal elements. This transformation of connection matrix and threshold vector change the shape of energy function but never change the energy values on the binary states of $\mathbf{x}$.

The network energy function of Eq. (1) can be minimized by using the following continuous state transition.

$$\frac{dx_i}{dt} = -\eta x_i(1 - x_i)[\mathbf{W}\mathbf{x} + \mathbf{th}]_i \quad (\eta > 0) \tag{4}$$

where $[\mathbf{a}]_i$ means the i-th component of the vector $\mathbf{a}$ in the brackets. $x_i(1 - x_i)$ is the factor for restricting the value x_i within the closed interval $[0, 1]$. This factor creates the non-linear state change of RNN. This state transition guarantees that the energy function of Eq.(1) always decreases independently of the diagonal elements of the matrix $\mathbf{W}$.

In the transformation of Eq.(2) and Eq.(3), it is important that this transformation never change the gradient vector $\dot{E}(\mathbf{o})$ at the center point $\mathbf{o} = \frac{1}{2}\mathbf{i} = (\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2})$ of the hypercube state space of RNN. This is proved by the following equations:

$$\dot{E}(\mathbf{o}) = \frac{dE(x)}{dx}\Big|_{x=o} \tag{5}$$

$$= \mathbf{W}\mathbf{o} + \mathbf{th} = \frac{1}{2}\mathbf{W}_0\mathbf{i} + \mathbf{th}_0$$

2.2. Expansion of Energy Function With Eigenvectors

The dynamical transition of RNNs can be roughly understood by introducing Eigenvalue analysis.

Letting the weight matrix $\mathbf{W}$ have the eigenvalue λ_i and the corresponding eigenvector $\phi_i, \{i = 1, \cdots, N\}$, an arbitrary state vector $\mathbf{x}$ can be decomposed into a linear combination of the eigenvectors as follows:

$$\mathbf{x} = \sum_i p_i\phi_i. \tag{6}$$

where p_i is a scaler variable which can change during the network state transitions. By substituting Eq.(6) into Eq.(1), the energy function can be expressed using eigenvalues and eigenvectors as follows:

$$E(\mathbf{x}) = \hat{E}(\mathbf{p}) = \frac{1}{2}\sum_i \lambda_i p_i^2 + \sum_i (\mathbf{th}^t\phi_i)p_i \tag{7}$$

$$= \frac{1}{2}\sum_i \lambda_i(p_i + \frac{\mathbf{th}^t\phi_i}{\lambda_i})^2 + Const.$$

Equation(7) means that the energy function can be also decomposed into the eigenvector components. Since the eigenvectors are linearly independent, the network behaviors can be estimated separately based on those components.

By multiplying the eigenvector ϕ_i to both side of Eq.(6), the variable p_i can be expressed by the inner product of the eigenvector ϕ_i and the state vector $\mathbf{x}$, that is $p_i = \phi_i^t\mathbf{x}$. From this, we can see, the terms $p_i + \frac{\mathbf{th}^t\phi_i}{\lambda_i}$ in Eq.(7) means the distance from a certain plane. Then we define the hyperplane Ω_i and the distance $d_i(\mathbf{x})$ from the hyperplane as follows:

$$Hyperplane : \Omega_i = \{\mathbf{x}|\phi_i^t\mathbf{x} + \frac{\mathbf{th}^t\phi_i}{\lambda_i} = 0\} \tag{8}$$

$$Distance : d_i(\mathbf{x}) = \phi_i^t\mathbf{x} + \frac{\mathbf{th}^t\phi_i}{\lambda_i} \tag{9}$$

By using the distance $d_i(\mathbf{x})$ of Eq.(9), the energy function of Eq.(7) becomes:

$$E(\mathbf{x}) = \frac{1}{2}\sum_i \lambda_i d_i(\mathbf{x})^2 + Const. \tag{10}$$

2.3. Eigenvalue Effects on Network State Transition

In order to decrease the energy function of Eq.(10), the network would change its state so that the distance $d_i(\mathbf{x})$ increases for a negative eigenvalue and decreases for a positive eigenvalue. If the eigenvalue is equal to 0, the hyperplane Ω_i does not exist. However, the network would change its state to decrease the energy function monotonously according to the first order term of Eq.(7). From this, we can classify the network behaviors into the following tree groups. Figure1 shows the image of the network behaviors of these groups.
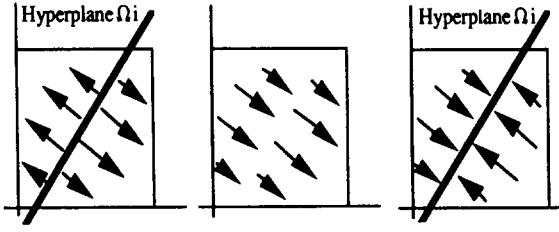
$$\Lambda_0 = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

$$\Phi_0 = [\phi_1, \phi_2, \ldots, \phi_N]$$

Then, the connection matrix of Eq.(2) can be expressed as follows;

$$\mathbf{W} = \mathbf{W}_0 + 2a\mathbf{I} \qquad (12)$$
$$= \Phi_0^t \Lambda_0 \Phi_0 + 2a\Phi_0^t \Phi_0$$
$$= \Phi_0^t (\Lambda_0 + 2a\mathbf{I})\Phi_0$$

Equation(12) means that the eigenvectors $[\phi_1, \phi_2, \ldots, \phi_N]$ of the connection matrix of Eq.(2) are invariant by uniformly changing diagonal elements. In addition, it means that the increment of the eigenvalues is as same as that of diagonal elements. Then, we can see that by increasing diagonal elements, negative eigenvalues decrease. As this result, a variety of convergence states of RNN decrease.

In addition, since all eigenvalues uniformly increase by increasing diagonal elements, patterns far from the eigenspace corresponding to higher eigenvalue are easy to lost the probability of convergence. Aiyer et al[8] suggested that the optimal solution of the network exists toward the direction of the eigenvector corresponding to minimum eigenvalue,(in their paper, described as maximum eigenvalue).

Then, the convergence to the optimal solution is thought to be more stable than that of another solution. Therefore,the reduction of convergence pattern by increasing diagonal elements improves the network performance for detecting the optimal solution.

Fig. 1: Network State Transition using Three Types of Eigenvalues

Negative eigenvalue: The network state moves so as to leave away from the hyperplane Ω_i. Then, the network hypercube are separated into two attractive regions by the hyperplane Ω_i.

Zero eigenvalue: The network state moves monotonously to the direction of decreasing the energy function. Then, Zero eigenvalue does not create the above separatability.

Positive eigenvalue: The network state moves toward the hyperplane Ω_i. Then, a positive eigenvalue also does not yield the separatability. However, if the hyperplane Ω_i does not include the vertex of the network hypercube, it may create a non-binary state convergence.

The network usually has many eigenvalues. The convergence results from the complex effects by these eigenvalues. However, since positive and zero eigenvalues do not create the separatability, the variety of the convergence is provided only by negative eigenvalues. Then, the network convergence patterns can be reduced by decreasing the number of negative eigenvalues.

3. Effects of Diagonal Elements on RNN Operation

3.1. Change of Eigenvalues

Let's consider the influence of diagonal elements on network dynamics using eigenvalue analysis. For this purpose, it is assumed that the connection matrix $\mathbf{W}_0$ of Eq.(2) can be transformed into the form with eigenvalue matrix Λ_0 and eigenvector matrix Φ_0 as follows:

$$\mathbf{W}_0 = \Phi_0^t \Lambda_0 \Phi_0 \qquad (11)$$

3.2. Movement of Eigenspace

Next, let's focus on the movement of eigenspace defined in Section 2. In order to make clear the movement of eigenspace, we pay attention to relation between the gradient vector $\dot{E}(o)$ at the center point and the normal vector $n_i(o)$ from the center point to eigenspace. The center point o is the center of state space of RNN and has the coordinate, $(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2})$. The relation between the normal vector $n_i(o)$ and eigenspace Ω_i is shown in Fig.2.

The normal vector $n_i(o)$ can be expressed such as the following equation, using the distance $d_i(o)$ from the center point o to a eigenspace Ω_i and eigenvector ϕ_i corresponding to the eigenspace.

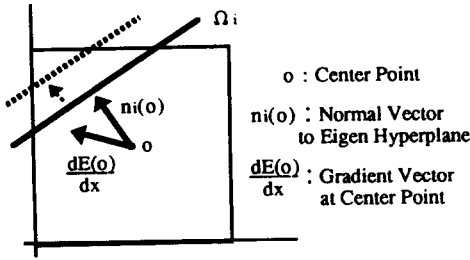$$n_i(o) = -d_i(o)\phi_i \qquad (13)$$

Fig. 2: Eigenspace Transition by Diagonal Elements

where the distance $d_i(\mathbf{o})$ is given by Eq.(9).

The distance $d_i(\mathbf{o})$ can be expressed using the product of the gradient vector $\dot{E}(\mathbf{o})$ and eigenvector ϕ_i, by reforming it as follows;

$$d_i(\mathbf{o}) = \frac{\lambda_i \phi^t \mathbf{o} + t h^t \phi_i}{\lambda_i} \qquad (14)$$

$$= \frac{(\mathbf{W}\mathbf{o} + t h)^t \phi_i}{\lambda_i} = \frac{\dot{E}(\mathbf{o})^t \phi_i}{\lambda_i}$$

Then, the normal vector $n_i(\mathbf{o})$ can be also written with the gradient vector $\dot{E}(\mathbf{o})$ and eigenvector ϕ_i, by substituting Eq.(14) to Eq.(13). Therefore, we can estimate the inner product of the gradient vector $\dot{E}(\mathbf{o})$ and the normal vector $n_i(\mathbf{o})$ as follows;

$$\dot{E}(\mathbf{o})^t n_i(\mathbf{o}) = -\frac{\{\dot{E}(\mathbf{o})^t \phi_i\}^2}{\lambda_i} \qquad (15)$$

The inner product of Eq.(15) means that the eigenspace Ω_i exists on higher energy side than the center point if its value is positive, and that Ω_i exists on lower energy side if it's negative. Since the numerator of Eq.(15) is always positive, the inner product of Eq.(15) has a positive value for a negative eigenvalue and a negative value for a positive eigenvalue.

As the diagonal elements of a connection weight matrix increase, all eigenvalues increase monotonously. When a negative eigenvalue approaches to 0, the corresponding eigenspace departs away from the center point toward the ascent direction of the energy slope. Vice versa, once a eigenvalue become greater than 0, the corresponding eigenspace approaches to the center point from the descent direction of the energy slope. The eigenspace corresponding to a negative eigenvalue acts as a divergence, while the eigenspace corresponding to a positive one acts as a convergence. Then, we can say, the movement of the eigenspace by increasing diagonal elements makes wider the attractive region to the descent direction of the center point of the state space of RNN.

It can be proved that the least energy binary state exists on the descent side of the center point $\mathbf{o}$. Then, we can say, the movement of eigenspace by diagonal elements contributes to the increase of the convergence probability to the optimal solution.

4. Simulation

4.1. 2-Dimensional Network

In order to make sure of the eigenvalue effects on the network dynamics, the behavior of the 2-dimensional network is simulated. Figure 3 shows the relation between the network dynamics and the eigenvalues. The lines in each graph shows the network state transition starting from random initial states. A circle is a convergence point.

The energy function is given by the following equation.

$$E(x_1, x_2) = 2x_1 x_2 - 1.6x_1 - 0.8x_2 \qquad (16)$$

$$+\frac{1}{2}a_1 x_1(1 - x_1) + \frac{1}{2}a_2 x_2(1 - x_2)$$

$$\qquad (17)$$

where,

$$\mathbf{W} = \begin{bmatrix} a_1 & 2 \\ 2 & a_2 \end{bmatrix}, \mathbf{th} = \left\{ \begin{array}{c} -\frac{1}{2}a_1 - 1.6 \\ -\frac{1}{2}a_2 - 0.8 \end{array} \right\}$$

where a_1 and a_2 are parameters to control the eigenvalues, which do not change the energy value at the vertices. The order of the energy levels at vertices is that $E(0,0) > E(1,1) > E(0,1) > E(1,0)$. Then, the optimal solution locates on the vertex$(1,0)$. The eigenvalues are shown under each graph.

In the 2-dimensional case, there are 2 eigenvalues in the connection weight matrix. If both are positive, the network converges to one point in the $[0,1]$ square region. This is a non-binary convergence. By decreasing the eigenvalues, the convergence point approaches to the edge. When one of the eigenvalues is equal to 0, the network converges only to the vertex $(1,0)$. By decreasing the eigenvalues more, the number of convergence states is increased gradually. When both eigenvalues are large negative, the number of convergence states become 4.

From this example, it is understood that the convergence probability to the optimal solution is improved by increasing eigenvalues. In this case, when one of the eigenvalue is 0 and another is positive, the convergence probability became best.

4.2. T.S.Problem

The method of controlling diagonal elements is applied to T.S.Problem. Fig. 4 shows the simulation results. The vertical axis of (b)(e) graphs show the amount of increments of diagonal elements. The

2645

(a) Diagonal elements: $a_1=a_2=$ +6.0
Eigenvalues = (2, 6)

(b) Diagonal elements: $a_1=a_2=$ +2.0
Eigenvalues = (0, 4)

(c) Eiagonal elements: $a_1=a_2=$ +0.0
Eigenvalues = (-2,2)

(d) Diagonal elements: $a_1=a_2=$ -3.0
Eigenvalues = (-4, 0)

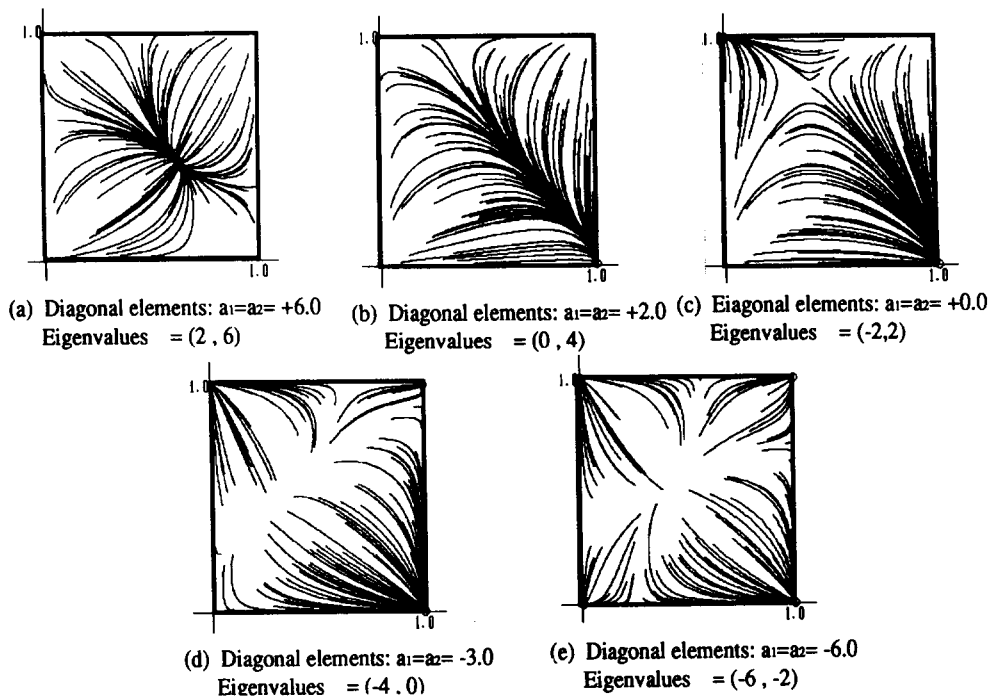(e) Diagonal elements: $a_1=a_2=$ -6.0
Eigenvalues = (-6, -2)

Fig. 3: Simulation of Network State Transition by Changing Eigenvalues

value 0 means that the maximum eigenvalue of connection matrix is 0, that is, all eigenvalues are not positive. Fig. 4(a) shows a pattern of city positions and the shortest path for traveling their cities. In this simulation, the condition of going back to a starting city is removed for increasing the variation of solutions.

In Fig. 4(b), the energy distribution of network convergence states is shown. The simulations were repeated 1000 times for each value of diagonal elements. The value increment of diagonal elements was 0.5. For the state change of RNN, the differential approximation of Eq.(4) was used. This state change operations were iterated until the network state didn't change. In this figure, the bold line shows the distribution of a binary convergence and the thin line shows the non-binary convergence. The alternative of binary or non-binary convergence was done by estimating whether all components of the network state vector are more than 0.9 or less than 0.1.

As diagonal elements increasing, the convergence changes from a binary state to a non-binary state. In addition, the energy of it decreases and the variance of the energy distribution becomes small.

Figure 4(c) shows the convergence probability to the optimal solution. The optimal solution was pre-searched by estimating all traveling patterns. The probability to the optimal solution increased until −8.0 of diagonal elements. However, over that

value, it rapidly fell down. This is by reason that the convergence to the optimal solution changed to a non-binary state.

Figure 4(d) shows the number of negative eigenvalues. As described in section 3, the number of negative eigenvalues decreases as increasing the diagonal elements. The decrease of the variance of convergence state is due to this decrease of negative eigenvalues.

Figure 4(e) shows the distances of eigenspaces from the center point. The sign of a distance is set such as the descent energy side of the center point is positive. From this graph, it is understood that all the distances first move to negative large value, and then, change from positive large value to 0. As described in section 2, this movement of eigenspace is one of the reasons of improvement of convergence performance of RNN with feedback loops.

From two simulations described in this section, it could be confirmed that the increase of diagonal elements can improve the convergence probability to the optimal solution. Then, in order to apply RNN to a combinatorial optimization problem, it is better to use RNN with feedback loops than that with no feedback loop. However, it has been also confirmed that if the diagonal elements are too large, the network would converge to a non-binary state, which is undesirable. Then the upper boundary is necessary for the diagonal elements.
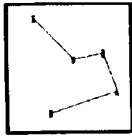
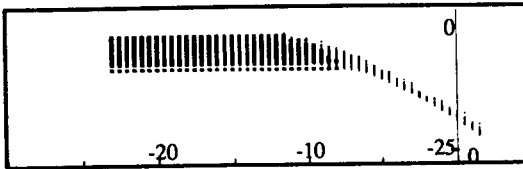Fig.4 (a) City Position

Fig.4 (b) Energy Distribution at Convergence States

Fig.4 (c) Convergence Probability to the Optimal Solution
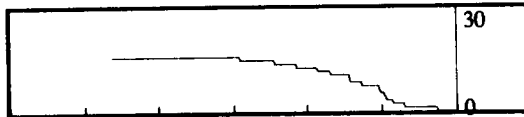

Fig.4 (d) Number of Negative Eigenvalues

Fig.4 (e) Distance from Eigenhyperplane

Fig. 4: Network Transition by Diagonal Elements

5. Conclusion

In this paper, we have proposed the analysis method using the eigenvalue and the hyperplane corresponding to the eigenvalue. Using this analyzing method, the transition of the number of negative eigenvalues and the movement of the eigenspaces have been discussed theoretically. As the results, it has been confirmed that both transitions by increasing diagonal elements contribute to improvement of the convergence probability to the optimal solutions. Moreover, this improvements have been confirmed by simulations using the 2-dimensional network and 5 cities T.S.problems.

However, the network with the nonzero diagonal elements may create non-binary solutions. In order to escape this undesirable solutions, the network needs the upper boundary for diagonal elements. It is difficult to decide this factor. Besides, it could not be proved theoretically whether the survival solutions on high diagonal elements include the optimal solution. These are next problem.

References

[1] Hopfield,J.J. and Tank,D.: *Neural Compu-* *tation of Decision in Optimization Problems.* Biol.Cybernetics, Vol.52, pp141-152,(1985)

[2] Abe,S., Kawakami,J. and Hirosawa,K.: *Solving Inequality Constrained Combinatorial Optimization Problems by the Hopfield Neural Network.* Neural Networks, Vol.5, pp663-670, (1992)

[3] Andrew H.Gee, Sreeram V.B.Aiyer and Richard W.Prager: *An Analytical Framework for Optimizing Neural Networks.* Neural Networks, Vol.6, pp79-97, (1993)

[4] Andrew H.Gee and Richard W.Prager: *Polyhedral Combinatorics and Neural Networks.* Neural Computation, Vol.6, pp161-180, (1994)

[5] Shigeo Abe, and Andrew H.Gee: *Global Convergence of the Hopfield Neural Network with Nonzero Diagonal Elements.* IEEE Trans. Circuits Syst., Vol.42, No.1, Jan(1995)

[6] Yoshinori Uesaka: "Mathematical Aspects of Neuro-Dynamics for Combinatorial Optimization", IEICE Trans., VOL.E74, No.6, pp1368-1372(JUNE1988)

[7] M Ohta, Y Anzai, S Yoneda and A Ogihara: "Theoretical Analysis of Neural Networks with Nonzero Diagonal Elements", IEICE Trans., VOL.E76, No.3, pp284-291(JUNE1988)

[8] Aiyer, S.V.B. Niranjan, M, and Frallside, F.: "A Theoretical Investigation into the Performance of the Hopfield Model", IEEE Trans.Neural Networks, vol.1, no.2, pp204-215, (Feb1990)