THE PUBLISHING HOUSE OF THE ROMANIAN ACADEMY
PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 14, Number 2/2013, pp. 95-100

ASYMPTOTIC EIGENVECTORS, TOPOLOGICAL PATTERNS AND RECURRENT NETWORKS

Alireza BAHRAINI

Sharif University of Technology, Dept. of Mathematical Sciences,
P.O.Box 11365-9415, Tehran, Iran
E-mail: bahraini@sharif.edu

The notions of asymptotic eigenvectors and asymptotic eigenvalues are defined. Based on these notions a special probability rule for pattern selection in a Hopfield type dynamics is introduced. The underlying network is considered to be a d-regular graph, where d is an integer denoting the number of nodes connected to each neuron. It is shown that as far as the degree d is less than a critical value d_c , the number of stored patterns with $m_\mu = O(1)$ can be much larger than that in a standard recurrent network with Bernouill random patterns. As observed in [4] the probability rule we study here turns out to be related to the spontaneous activity of the network. So our result might be an evidence for the idea that some spontaneous activities intend to reorganize information for improving the capacity .

Keywords: signal-noise analysis, asymptotic eigenvalue, asymptotic eigenvector, Hopfield model.

1. INTRODUCTION

Let $A=[a_{ij}]_{N\times N}$ be the adjacency matrix of a simple graph representing the architecture of a neural network consisting of N neurons. This means that for $1\le i,j\le N, a_{ij}=1$, if the neurons i and j are interconnected and $a_{ij}=0$ if they are not so. Assume that the microscopic state of the network is defined by the N-neuron state vector $\sigma=(\sigma_1,...,\sigma_N)\in\{\pm1\}^N$ where $\sigma_i=1$ if neuron i fires and $\sigma_i=-1$ if it is at rest. Let the network evolves according to a standard Glauber-type dynamics. The local field at neuron i is given by $h_i=\sum_j J_{ij}\sigma_j$, and the interactions matrix $[J_{ij}]_{N\times N}$ is determined through the well-known Hebb rule $J_{ij}=\frac{a_{ij}}{N}\sum_{\mu=1}^{p}\xi_i^\mu\xi_j^\mu$, where $\xi^\mu=(\xi_1^\mu,...,\xi_N^\mu)\in\{\pm1\}^N, \mu=1,...,p$, denote the p patterns to be stored in the system. In a deterministic parallel dynamics, the evolution of the system is described as usual by the following rule, $\sigma_i(t+1)=\operatorname{sgn}(h_i(\sigma(t)))$. One of the most studied probability rules for pattern selection in a recurrent network is the following well-known Bernouill probability rule $\Pr(\xi_i^\mu=1)=\frac{1}{2}(a+1)$, for $\mu=1,...,p$, and for, $i=1,...,p$, where $0\le a\le 1$ is a real positive number. The study of dynamics of networks in which more complicated probability rules for pattern selection are applied is not in general so simple and we do not yet know any systematic work in this direction.

B.B. Averbeck, P.E.Latham and A. Pouget in [2] have provided qualitative answers to the following question: Does adding correlations to a population of neurons without modifying single neuron responses (so that the correlated and uncorrelated populations would be indistinguishable on the basis of single neuron recordings) increase or decrease the amount of information in the population? What we present in this note is part of research carried out in the direction of the above fundamental question. There exist other reasons for our interest to this problem. In realistic complex networks the observed patterns seem to be more complicated than the patterns selected according to the above simple Bernouoill probability rule. In fact according to experimental data for example in neocortical areas the observed patterns seem to have correlations with the topography of the network. (see e.g. [3, 4, 5]) Our aim in this note is to show that for certain classes of large graphs representing the connectivity between neurons in a recurrent neural network there exits a collection of patterns which are more compatible with the architecture of the network and are best suited to be stored in it. More precisely we will introduce a collection of asymptotic eigenvectors $\{V^{\mu}\}_{\mu}$, with (positive) asymptotic eigenvalues $\{\lambda_{\mu}\}_{\mu}$ associated to the underlying topology, and we suppose that patterns are selected among this special collection according to the following probability rule:

$$\Pr(\xi = V^{\nu}) = \frac{\lambda_{\nu}}{\sum_{\mu} \lambda_{\mu}}, \qquad \nu = 1, ..., N. \tag{1}$$

This probability rule for pattern selection could be justified for example by the observation in [4], stating that in a spontaneous activity of a network, eigenvectors of the connectivity matrix of that network appear with a probability proportional to their eigenvalues. We have also provided in [1] a justification on how similar probability rules could appear as rules after phase transition in a Gibbs type topological probability rule of pattern selection. In order to clarify the idea we start by describing a simple example.

2. RING MODEL

Consider a network on a ring with N nodes. The connectivity matrix $A = [a_{ij}]$ of such a network is given by:

$$a_{ij} = \begin{cases} 1 & \text{if} \quad j = i+1 \mod N \\ 0 & \text{otherwise} \end{cases}. \tag{2}$$

Among all the patterns $\xi = (\xi_1, ..., \xi_N) \in \{\pm 1\}^N$, we consider the class of -2-block patterns consisting of those patterns in which all the +1's are set next to each other. We claim now that these 2-block patterns being more compatible with the geometry of the network, are in some sense more appropriate to be stored in it. To show this we manually select the following collection of $[\frac{N}{2}]$, 2-block patterns:

$$\xi_i^k = \begin{cases} +1 & \text{if} \ 1 \le i \le 2k \\ -1 & \text{otherwise} \end{cases}, \tag{3}$$

where $0 \le 2k \le N$. According to Hebb rule, the associated interaction weights J_{ij} are given by:

$$J_{2k,2k+1} = \frac{m-2}{2m}, \quad \text{for} \quad k = 1, ..., m-1$$
$$J_{2k-1,2k} = \frac{1}{2} \quad \text{for} \quad k = 1, ..., m \tag{4}$$

for $N = 2m$; and by:

$$J_{2k,2k+1} = \frac{m-2}{2m+1}, \quad \text{for} \quad k = 1, ..., m$$
$$J_{2k-1,2k} = \frac{2m}{2m+1}, \quad \text{for} \quad k = 1, ..., m \tag{5}$$
$$J_{2m+1,1} = -\frac{m}{2m+1}$$

for $N = 2m+1$.

Let also R_{2l}, for $l=1,...,[\frac{N}{2}]$, be an operator acting on the space of patterns as follows: $R_{2l}(\xi)$ is a pattern which is obtained from $\xi=(\xi_{1,},...,\xi_N)\in\{\pm1\}^N$ by a cyclic permutation of the form:

$$R_{2l}(\xi)=(\xi_{2l+1},...,\xi_N,\xi_1,...,\xi_{2l}). \tag{6}$$

One can easily verify that not only the $[\frac{N}{2}]$ patterns $\{\xi^{2k}\}$, $k=1,...,\left[\frac{N}{2}\right]$, introduced by (3) are stored in the network but also all the patterns of the form $R_{2l}(\xi^{2k})$, for $k,l=1,...,\left[\frac{N}{2}\right]$, can be retrieved in a Hopfield-type dynamics. Consequently at least $[N/2]^2$, 2-block patterns can be stored in a Ring model.
In order to generalize the phenomenon observed in this special example we will introduce the notions of asymptotic eigenvectors and asymptotic eigenvalues in the next section.

3. ASYMPTOTIC EIGENVECTORS

Let $A=[a_{ij}]$ be an $N\times N$ matrix with $N>>0$. We first introduce a definition:

Definition 1. A real vector $V\in\mathbb{R}^N$ is called an asymptotic eigenvector if there exists a real number $\lambda\in\mathbb{R}$ such that:

$$(A-\lambda I)\frac{NV}{|V|}=O\left(\frac{1}{N}\right), \tag{7}$$

where I is the identity matrix of dimension N. We call λ an asymptotic eigenvalue.

Definition 2. A graph G with its connectivity matrix A is called a literate graph if there exits real parameters κ and ϑ such that the operator

$$\Delta=\kappa I-\vartheta A \tag{8}$$

is positive definite and there exists a collection $V^\mu\in\{\pm1\}^N$, $\mu=1,...,M$, of asymptotic eigenvectors with asymptotic eigenvalues $\{\lambda^\mu\}_{\mu=1}^M$ associated to Δ such that the following spectral relation holds:

$$\sum_{\mu=1}^{N} a_{ij}\frac{\lambda_\mu}{\sum_k \lambda_k}V_i^\mu V_j^\mu=r\Delta_{ij}a_{ij}+O\left(\frac{1}{N}\right), \tag{9}$$

where $r\in\mathbb{R}$ is a real constant.

Example (Ring model.) Let A be the connectivity matrix of a ring of N nodes and set $\Delta:=I+A$. Let $1<q<N$ and let V^μ, for $\mu=1,...,N$ be the set of 2-block patterns with exactly q coefficients equal to 1. It can be easily verified that $\{V^\mu\}_{\mu=1}^N$ constitute a set of asymptotic eigenvectors of Δ with a common eigenvalue $\lambda_\mu=3$, for $\mu=1,...,N$. The collection $\{V^\mu\}_{\mu=1}^N$ satisfy (9) with $r=1$.

4. MAIN THEOREM

Let A be the connectivity matrix of a *literate graph* G of N nodes and let the degree of each node be equal to a given integer *d*. Suppose that $V^\mu\in\{\pm1\}^N$, for $\mu=1,...,N$, constitute a collection of asymptotic eigenvectors for A with positive asymptotic eigenvalues $\lambda_\mu \in \Re$, $\mu = 1,...,N$ satisfying (9). Assume that p patterns are selected among the eigenvectors $\{V^\mu\}_{\mu=1}^N$, according to the probability rule (1) and synaptic weights have been modified according to the well-known Hebb rule as described in the introduction. Consider a Hopfield type network on G with its standard Glauber-type dynamics. Then we have the following theorem :

THEOREM 1. *If* $\lambda = \min_\mu \{\lambda_\mu\}$, *then there exists a constant* β *such that if* d *satisfies* $d \le d_c = r(\lambda - \kappa)\dfrac{\sqrt{p}}{2\beta \ln N}$ *, then all the patterns* V^μ *for which* $0 < \lambda_\mu - \kappa = O(1)$ *are stored in the network.*

Proof. Let p patterns $\xi^\mu, \mu = 1,..., p,$ be selected through the probability rule (9). For simplicity of notation suppose that $\xi^1 = V^1$. We would like to proceed a signal-noise analysis. The signal part of the local field $h_i = \sum_j J_{ij} V_j^1$ is equal to :

$$S_i = \frac{1}{N}(\sum_j a_{ij} (V_j^1)^2) V_i^1 = \frac{d}{N} V_i^1. \tag{10}$$

The noise term comes from

$$R_i = \frac{1}{N} \sum_{j,\mu \ne 1} a_{ij} \xi_i^\mu \xi_j^\mu V_j^1. \tag{11}$$

Using (7) for $\mu = 1,\dots, p$, we can write:

$$< \xi_i^\mu \xi_j^\mu > = \sum_\mu a_{ij} \frac{\lambda_\mu}{\sum_k \lambda_k} V_i^\mu V_j^\mu = r a_{ij} \Delta_{ij} + O\left(\frac{1}{N}\right). \tag{12}$$

Thus we reach to:

$$< R_i > = \frac{p}{N} \sum_j r a_{ij} \Delta_{ij} V_j^1 + O\left(\frac{p}{N^2}\right) \tag{13}$$

by definition:

$$\Delta_{ij} = \kappa \delta_{ij} - \vartheta a_{ij}, \tag{14}$$

so:

$$\sum_j a_{ij} \Delta_{ij} V_j^1 = \sum_j a_{ij} (\kappa \delta_{ij} - \vartheta a_{ij}) V_j^1 = -\vartheta \sum_j a_{ij} V_j^1. \tag{15}$$

Using the relation $\Delta V^1 = \lambda_1 V^1 + O\left(\frac{1}{N}\right)$ and (14) we obtain:

$$\vartheta \sum_j a_{ij} V_j^1 = (\kappa - \lambda_1) V_i^1 + O\left(\frac{1}{N}\right), \tag{16}$$

so we have:

$$< R_i > = \frac{pr}{N} (\lambda_1 - \kappa) V_i^1 + O\left(\frac{p}{N^2}\right). \tag{17}$$

As can be seen if $\lambda_1 \geq \kappa$ this term has a positive effect on the signal term S_i and serves for stabilizing the corresponding pattern V^1. Consequently the total signal term can be redefined as:

$$S_i + \langle R_i \rangle = \frac{d + rp(\lambda_1 - \kappa)}{N} V_i^1. \quad (18)$$

In order to estimate the variance of R_i we write:

$$R_i = \frac{1}{N} \sum_{\mu} \xi_i^{\mu} R_i^{\mu}, \quad (19)$$

where

$$R_i^{\mu} := \sum_j a_{ij} \xi_j^{\mu} V_j^1. \quad (20)$$

This is in general something of $O(d)$ and so:

$$\mathrm{Var}(R_i) = O\left(\frac{pd^2}{N^2}\right) = \beta \frac{pd^2}{N^2}, \quad (21)$$

for some constant β. The probability for $\sigma_i(0) = V_i^1$ to be stable is equal to the probability that,

$$\Pr(V_i^1 h_i) > 0. \quad (22)$$

On the other hand the signal to noise ratio is equal to :

$$\frac{\sqrt{\mathrm{Var}(V_i^1 R_i)}}{\langle (S_i + R_i) V_i^1 \rangle} \leq \frac{\beta \sqrt{p} d}{(d + rp(\lambda_1 - \kappa))}, \quad (23)$$

for some constant β. Now if $p >> d$ the dominant term of the right hand side is given by:

$$\frac{\beta \sqrt{p} d}{rp(\lambda_1 - \kappa)} = \frac{\beta d}{r \sqrt{p} (\lambda_1 - \kappa)}. \quad (24)$$

This shows that if $\lambda_1 - \kappa = O(1)$, then there exists a critical value d_c for the degree of each node in the network, such that as long as $d \leq d_c$ the pattern V^1 remains stable with probability one at almost all the nodes. This critical value is easily seen to be equal to:

$$d_c = r(\lambda_1 - \kappa) \frac{\sqrt{p}}{2\beta \ln N}. \quad (25)$$

In the case where $p = O(N)$, we get $d_c = O(r(\lambda_1 - \kappa)\frac{\sqrt{N}}{2 \ln N})$ and the theorem is established.

ACKNOWLEDGEMENTS

The author thanks the research council of Sharif University of Technology for support.

REFERENCES

1. A. BAHRAINI, A. ABBASSIAN, *Topological Pattern Selection in Recurrent Networks*, Preprint.
2. AVERBECK,B.B., LATHAM, P.E., POUGET, A., *Neural correlations, population coding and computation*, Nature Reviews, Neuroscience, Vol 7, May 2006, pp. 358–366.
3. FELLEMAN, D.J. VAN ESSEN, D.C., *Distributed hierarchical processing in the primate cerebral cortex*, 1991, Cereb. Cortex 1 1.
4. GALAN R.F., *On How Network Architecture Determines the Dominant Patterns of Spontaneous Neural Activity*, PLoS ONE **3**, (5); e2148; doi:10.1371/ journal.pone.0002148, 2008.
5. ROUDI, Y., TREVES, A., *An associative network with spatially organized connectivity*. J. Stat. Mech., P07010, 2004.

Received January 3, 2012