УДК 519.71, 612.82

STOCHASTIC DYNAMICS OF NEURAL NETWORKS AFTER LEARNING

V.V. KAPELKO, A.D. LINKEVICH
(Polotsk State University)

We study probabilistic synchronous dynamics of Little-Hopfield neural networks with asymmetric interneuronal synaptic connections adjusted in accordance with a learning rule given in [3]. Types of behaviour of such systems are analysed in dependence of a vector-parameter characterizing degrees of freedom in determining synaptic couplings. We have found that in the case of small level of noise networks can store memorized patterns but as an amount of noise is increased behaviour becomes more complex and, beginning with some critical value the motion seems to be chaotic.

In recent years a significant progress have been achieved in study of information processing performed by neural networks. A considerable role in this advance belongs to the Hopfield paper [1]. However this model is based on a number of simplifications produced features which are biologically not plausible. So, various sources of synaptic and threshold noise arise in real nervous systems which lead rather to a stochastic neurodynamics instead of the deterministic one; interneuronal couplings (synaptic efficacies) are asymmetric, if not unidirectional; connectivity of neurons is high but not complete; all synapses ,which connect a given neuron to others, are presumably either excitatory or inhibitory; etc.

Attempts are undertaken to approach to a more adequate description of neurosystems and many interesting results have been obtained along this line (see, for a review [2]). So, neural networks with asymmetric synaptic connections display a diversity of types of dynamical behaviour including relaxation towards fixed points and limit cycles, chaotic motions, temporal association of patterns and so on. More effective learning algorithms are designed which should adjust the synaptic efficacies so as to ensure abilities of the networks required for information processing.

In this paper we study probabilistic synchronous dynamics of binary neural networks with asymmetric synaptic couplings adjusted in accordance with a learning rule given in [3]. We have obtained that in the case of small level of the threshold noise accounted by an effective temperature the networks indeed do store memorized patterns. As an amount of the noise is increased, behaviour of the systems become more complex and, begining with some critical value, the motion seems to be chaotic.

Consider a network composed of N neurons whose states are described by binary variables so that $s_i = +1$ means that neuron i is firing and $s_i = -1$ corresponds to the case when neuron i is quiescent. Suppose that the probability for neuron i be in state υ_i at time moment $t+1$ is given by the equation [4]

$$P_i(\sigma_i) = \frac{1}{2}\left\{1 + \sigma_i \cdot \tanh\left(\beta h_i\left[\vec{s}(t)\right]\right)\right\}. \tag{1}$$

Here the local field

$$h_i\left[\vec{s}(t)\right] = \sum_{j=1}^{N} J_{ij} s_j(t) - T_i \tag{2}$$

is expressed through the synaptic efficacies J_{ij} and the thresholds T_i (hereafter taken to be zero for simplicity). The effective temperature $1/\beta$ takes into account the threshold noise.

The synaptic matrix J_{ij} should be adjusted while learning in such a way that a prescribed set of memorized patterns $\{\xi_i^\mu\}$, $i = \overline{1, N}$, $\mu = \overline{1, p}$ be fixed points of the network dynamics if the noise is small enough $(1/\beta \to 0)$. We adopte the following prescription [3]

$$J_{ij} = \delta_{ij} + C_i \theta_j, \tag{3}$$

where vector $\vec{\theta} = (\theta_1, \dots \theta_N)$ is constructed so as to be orthogonal to all memorized patterns, i.e. $\left(\vec{\theta} \cdot \xi^\mu\right) = 0$ for all μ, and $\vec{C} = (C_1, \dots C_N)$ is an arbitrary vector. Choice of vector $\vec{C}$ influences strictly on behaviour of the system. If we put $\vec{C}$ to be proportional to $\vec{\theta}$ than eq. (3) yields the symmetric matrix J_{ij} which is close to the synaptic matrices produced by the so-called pseudo-inverse learning rules [5] because $J\vec{\xi}^{\mu} = \vec{\xi}^{\mu}$ due to eq. (3) (for arbitrary $\vec{C}$). In computer simulations a considerable difference was not observed between both cases. Further we are concentrated on the case of asymmetric matrices J_{ij}. To analyse the network behaviour, it is useful to consider the quantity $h_i s_i$ which, in view of eqs. (2), (3), can be expressed as

$$h_i s_i = 1 + kC_i \cdot s_i = 1 + k|C_i| \operatorname{sgn} C_i \cdot s_i \tag{4}$$

where $k = \vec{\theta} \cdot \vec{s}$. If we denote

$$b = \inf_{s_i \ldots s_n} |k|, \qquad B = \sup_{s_i \ldots s_n} |k| = \sum |\theta_j|, \tag{5}$$

then it is obvious that (for $k \neq 0$) $b \leq |k| \leq B$. From eq. (1) we can conclude that if $h_i s_i > 0$ for some state $\vec{s} = (s_1, \ldots s_N)$ then it is more probable for neuron i to keep its state in next time moment. The greater the value of $h_i s_i$ the greater this probability. Conversely, if $h_i s_i < 0$ then the change of state s_i is more probable.

First consider the case when $|C_i| \cdot B < 1$ for some i. Then

$$|C_i| \cdot |k| < 1$$

and, in view of eq. (4), we find that $h_i s_i > 0$ for an arbitrary state $\vec{s}$, i.e. neuron i rather keeps its state s_i intact. If the condition aforesaid is satisfied for all neurons then every state $\vec{s}$ of the network is a fixed point of the deterministic dynamics but not only the memorized patterns. This case is not interesting with the standpoint of associative memory and further it will be outside our cosideration.

Now let the condition $|C_i| \cdot b > 1$ be held for some neuron i. Then we have that $|C_i| \cdot |k| > 1$ and therefore the sign of the quantity $h_i s_i$ coincides with the sign of the product $(k s_i \operatorname{sgn} C_i)$. If $k > 0$ then neuron i tends to go into the state $s_i^* = \operatorname{sgn} C_i$ and for $k < 0$ the most probable transition is into the state $-s_i^*$.

The most interesting types of the behaviour occur when the intermediate conditions

$$1/B < |C_i| < 1/b \tag{6}$$

are satisfied. Results relevant to this case will be given below while here it is expedient to note that eqs. (6) restrict the absolute values of the components C_i but tell nothing about their signs. On the other hand, choice of the values $\operatorname{sgn} C_i$ influences highly on the network behaviour while it gets in the state s_i^* or $-s_i^*$.

Indeed, let us assume that the network is in the state $\vec{s}^*$ and consider possible variants of its further behaviour which depend on the value of the quantity

$$k^* \equiv \vec{\theta} \cdot \vec{s}^* = \sum \theta_i \cdot \operatorname{sgn} C_i \tag{7}$$

If k^* is positive then we find from eq. (4) that $h_i\left[\vec{s}^*\right] \cdot s_i^* > 1$ while for every memorized pattern $h_i\left[\vec{\xi}^{\mu}\right] \cdot \xi_i^{\mu} = 1$. This means that the state $\vec{s}^*$ is a fixed point (in the deterministic case) and besides, for the stochastic dynamics, the probability of keeping this state is greater than the probabilities of keeping the states $\vec{\xi}^{\mu}$. Simulations support that this makes worse retrieval of memorized patterns. On the other hand, if the quantaty k^* has such a value that $h_i\left[\vec{s}^*\right] \cdot s_i^* < 0$ than the cycle with period-2 appears between the states $\vec{s}^*$ and $-\vec{s}^*$. A greater interest is attracted to the intermediate region when

$$-1\big/\min_i |C_i| < k^* < -1\big/\max_i |C_i|. \tag{8}$$

Here the most probable event is such that the network turns from the state $\vec{s}^*$ into a state which is different both from $\vec{s}^*$ and $-\vec{s}^*$.

We conducted computer simulations for the case when vector $\vec{C}$ satisfies the conditions (6) and (8). The typical number of neurons was taken to be $N = 12$, the number of memorized (random) patterns was $p = 4$ (we were mainly concetrated on networks of small sizes since if such a system possess some complex behaviour then more large networks can obviously share the types of motion observed).

One of the most important question in exploring network behaviour is about types of motion after transient processes. So, if a network keeps its state after reaching a memorized pattern then this means retrieval of this pattern. But if, instead, a network makes a transition from state $\vec{\xi}^{\mu}$ to some unstable orbit, then such a motion is outside autoassociative memory. To investigate this property, we put the initial state of a network into memorized pattern $\vec{\xi}^1$ and consider further behaviour. We calculated time evolution of the overlap

$$m^1(t) = \frac{1}{N}\sum_{i=1}^{N} s_i(t)\xi_i^1$$

between the network state and the pattern. We find also the Fourier spectrum and autocorrelation function computed for the overlap $m^1(t)$ in the case of $\beta = 1$ and $\beta = 3$. We observed that in the case of small β (large noise) time evolution of the overlap is irregular (broad power spectrum and fast decreasing the autocorrelation function). As β is increased, the network tends to spend more time in fixed points, exhibiting sometimes oscillatory bursts. For sufficiently large β $(\beta > 5)$ the network reaches a memorized pattern and stays in this state for very long time (of order thousands iterations). This case corresponds to retrieval of the stored information. (We analyzed also overlap histograms obtained from simulations of networks with the same initial conditions and synaptic matrices).

REFERENCES

1. Hopfield J.J. Proc. Natl. Acad. Sci. USA, 1982, 79, 2554.
2. Linkevich A.D. Proc. Seminar "Nonlinear Phenomena in Complex Systems", Polatsk, February 17 – 20, 1992. – P. 21.
3. Linkevich, A.D. Learning neural networks with the aid of projection techniques, Proceedings of the Second Seminar "Nonlinear Phenomena in Complex Systems", Polotsk, February 15 – 17, 1993. – P. 373.
4. Little W.A. Math. Biosci., 1974, 19, 281.
5. Kohonen, T. Associative Memory / T. Kohonen. – Berlin: Springer, 1978; Personnaz L., Guyon I., Dreyfus G.J. Physique Lett., 1985, 46, L359; Diederich S., Opper M. Phys. Rev. Lett., 1987, 58, 949; Berryman K.W. Inchiosa M.E., Jaffe A.M., Janowsky S.A. J. Phys. A, 1990, 23, L223; Linkevich A.D. J. Phys. A, 1992, 25, 4139.

Поступила 01.09.2011