Neurocomputing 141 (2014) 188–193

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A class of stochastic Hopfield neural networks with expectations in coefficients

Yangzi Hu

School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, PR China

ARTICLE INFO

Article history:
Received 1 July 2013
Received in revised form
15 February 2014
Accepted 23 March 2014
Communicated by W. Lu
Available online 12 April 2014

Keywords:
Stochastic Hopfield neural network
Time-varying delay
Exponential stability
Expectation

ABSTRACT

This paper considers a class of stochastic Hopfield type neural networks with time-varying delays. Based on the architecture of neural networks, expectations are introduced into coefficients of the model we considered. By applying the *M*-matrix technique, the mean square exponential stability is studied. An example is provided to illustrate the effectiveness of our results.

1. Introduction

The neural network proposed by Hopfield in 1980s can be described by an ordinary differential equation of the form

$$C_i\frac{dx_i(t)}{dt} = -\frac{1}{R_i}x_i(t) + \sum_{j=1}^{n} T_{ij}g_j(x_j(t)) + I_i, \quad 1 \le i \le n$$

where $x_i(t)$ denotes the voltage on the input of the ith neuron; C_i denotes the neuron input electric capacity; R_i denotes the neuron transmission resistance; T_{ij} denotes the interconnected synaptic character between neurons and I_i denotes the external input electric current to the ith neuron.

After decades, research on Hopfield neural networks has advanced greatly. Taking stochastic disturbances and time delays into consideration are two important aspects of improvements in Hopfield neural networks. As pointed out by Haykin [1] that in real nervous systems, synaptic transmission is a noisy process brought on by random fluctuations from the release of neurotransmitters and other probabilistic causes. On the other hand, time delays are unavoidable in hardware implementation, due to the finite switching speed of amplifiers or finite speed of information processing, and the existence of time delays may lead to oscillation, divergence, and even instability in network systems [2]. It is therefore reasonable to consider stochastic neural networks with time delays. Some results related to this problem have been published, for example, see [3–14].

First, we give a typical form of stochastic Hopfield neural networks with delays, which is also a special case of the neural network models considered in our previous work[15]:

$$dx(t) = [-Bx(t) + Ag(y(t))]\,dt + \sigma(t, x(t), y(t))\,dw(t), \tag{1.1}$$

where $B = \text{diag}(b_1, b_2, \ldots, b_n)$ with $b_i > 0$ $(1 \le i \le n)$ and $A \in \mathbb{R}^{n\times n}$;

$$y(t) = (y_1(t), y_2(t), \ldots, y_n(t))^{\mathrm{T}}, \quad y_i(t) = x_i(t-\delta_i(t)) \ (1 \le i \le n) \tag{1.2}$$

where $\delta_i(t)$ $(1 \le i \le n)$ are variable delays. $w(t)$ is an m-dimensional Brownian motion. In the following, we will improve this model by adding expectations.

Psychophysical experiments and analysis of the data given by a linear neuronal model showed that it is reasonable to investigate mechanism of expectation on probabilistic stimuli [16]. The results suggest that the expectation is held by synaptic weight of the neuron and the neural mechanism controlling saccade based on expectation consists of two pathways working in parallel. One facilitates saccades based on expectation. The other executes saccades in response to the saccade signal. Also, [17] pointed out that a neuron is connected to one another. There is a real number associated with each connection, which is called the weight of the connection. A neuron in the output layer determines its activity by a two step procedure. First, it computes the total weighted input x_i, using the formula: $x_i = \sum_{j=1}^{n} y_j p_{ji}$, where y_j is the activity level of the ith neuron in the previous layer and p_{ji} is the weight of the connection between the ith and the jth neuron. Hence, we may introduce expectation $\mathbb{E}x_i(t)$ into model (1.1) to describe this process. Next, the neuron calculates the activity y_i by using the sigmoid function of the total weighted input, and then the activities of all output units can be determined. The second step

E-mail address: huyangzi03@hotmail.com

http://dx.doi.org/10.1016/j.neucom.2014.03.021
0925-2312/

is described by the term $Ag(y(t))$. However, we know that Hopfield neural networks are recurrent, that is, the output of any layer affects that same layer, so the output perhaps becomes the input. Hence, we take its weight $\mathbb{E}y(t)$ into consideration. Then we investigate the following stochastic neural networks in this paper:

$$dx(t) = [-Bx(t) + Ag(y(t)) + R\mathbb{E}x(t) + S\mathbb{E}y(t)]\,dt + \sigma(t, x(t), y(t), \mathbb{E}x(t), \mathbb{E}y(t))\,dw(t). \tag{1.3}$$

Let $x(t)$ $(-\tau \le t < \infty)$ be a solution to Eq. (1.3) with $\tau = \max_i \delta_i(0)$. ξ is the initial data of $x(t)$ which means that $x(\theta) = \xi(\theta)$ for any $\theta \in [-\tau, 0]$. We denote $x(t, \xi)$ the unique solution of Eq. (1.3) with initial data ξ. Kloeden and Lorenz [18,19] called equations which are similar to Eq. (1.3) as a class of mean-field stochastic differential where other sample paths influence the evolution of a sample path of the solution.

Many applications of neural networks are dependent on the stability. Stability is therefore an important topic when designing neural networks. In this paper, we focus on the mean square exponential stability of model (1.3), namely,

$$\limsup_{t\to\infty} t^{-1} \ln \mathbb{E}|x(t)|^2 \le -\gamma, \tag{1.4}$$

where γ is a positive constant independent of the initial data. Eq. (1.4) shows that when $t \to \infty$, $\mathbb{E}x_i^2(t)(1 \le i \le n)$ exponentially decays to zero. Also, we wonder that how the two terms $R\mathbb{E}x(t) + S\mathbb{E}y(t)$ affect the stability of model (1.3). Roughly stated, our answer is that: $R\mathbb{E}x(t) + S\mathbb{E}y(t)$ tends to damage the stability of neural networks. Hence, we can say that to figure out instability problems, model (1.3) is much better than model (1.1). This also means that stability results given by model (1.1) may not be reliable. Model (1.3) perhaps gives more correctly prediction. Our example in Section 4 will illustrate this phenomenon intuitively.

Methodologically, the semi-martingale convergence theorem cannot be used anymore, because model (1.3) involves expectations $\mathbb{E}x(t)$ and $\mathbb{E}y(t)$. Fortunately, we find a method to deal with the system involving expectations and then obtain our desired results.

2. Preliminaries

For any $x \in \mathbb{R}^n$ and $\mathbb{R}^n$-valued function f, we always assume that

$$x = (x_1, x_2, \ldots, x_n)^T, \quad f = (f_1, f_2, \ldots, f_n)^T; \quad \mathrm{diag}(x) = \mathrm{diag}(x_i) = \mathrm{diag}(x_1, x_2, \ldots, x_n).$$

Let $\tau = \max_i \delta_i(0)$. Denote that function space $C = C([-\tau, 0], \mathbb{R}^n)$ with the supremum norm: $\|\xi\| = \sup_{-\tau\le\theta\le 0}|\xi(\theta)|(\xi \in C)$. Let $|\cdot|$ denote the Euclidean norm of vectors or the trace norm of matrices. Denote that $\Delta_i(t) = t - \delta_i(t)$ $(1 \le i \le n)$. Suppose $\delta_i(t) \in C^1(\mathbb{R}_+)$,

$$0 \le \delta_i(t) \le b < \infty \quad (t \ge 0) \tag{2.1}$$

$$\eta_i := \inf_{t\ge 0} \Delta_i'(t) > 0. \tag{2.2}$$

Eq. (2.2) implies that $\eta_i \le 1$ and $\lim_{t\to\infty}\Delta_i(t) = \infty$. Obviously, $\Delta_i(t)$ is strictly increasing on $[0, \infty)$, and its inverse function $\Delta_i^{-1}(s)$ is defined on $[-\delta_i(0), \infty)$, and has property that

$$[\Delta_i^{-1}(s)]' = 1/\Delta_i'(t) \le \eta_i^{-1} \quad (s = \Delta_i(t), t \ge 0). \tag{2.3}$$

Assume that both

$$g(y) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

and

$$\sigma(t, x, y, u, v) : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times m}$$

are Borel measurable functions and satisfy the local Lipschitz condition. Write that

$$f(x, y, u, v) = -Bx + Ag(y) + Ru + Sv, \tag{2.4}$$

$$F(t) = f(x(t), y(t), \mathbb{E}x(t), \mathbb{E}y(t)), \quad \Sigma(t) = \sigma(t, x(t), y(t), \mathbb{E}x(t), \mathbb{E}y(t)). \tag{2.5}$$

Using (2.5), Eq. (1.3) can be written as

$$dx(t) = F(t)\,dt + \Sigma(t)\,dw(t)$$

or

$$x(t) = x(0) + \int_0^t F(s)\,ds + \int_0^t \Sigma(s)\,dw(s). \tag{2.6}$$

For any given $V(x) \in C^2(\mathbb{R}^n)$, $t \ge 0$ and $x, y, u, v \in \mathbb{R}^n$, define that

$$\mathcal{L}V(t, x, y, u, v) = V_x(x)f(x, y, u, v) + \tfrac{1}{2}\mathrm{tr}[\sigma^T(t, x, y, u, v)V_{xx}(x)\sigma(t, x, y, u, v)], \tag{2.7}$$

where f is given by (2.4). If $x(t)$ is a solution to Eq. (1.3), then applying the Itô formula and (2.7) we have that

$$dV(x(t)) = LV(x(t)) + V_x(x(t))\Sigma(t)\,dw(t), \tag{2.8}$$

where $LV(x(t)) = \mathcal{L}V(t, x(t), y(t), \mathbb{E}x(t), \mathbb{E}y(t))$ with $y(t)$ in (1.2) and $\Sigma(t)$ in (2.5).

The following lemma is important whose similar form can be found in [20], so we omit the proof here.

Lemma 2.1. *Let $x(t)$ $(-\tau \le t < \infty)$ be a solution to Eq. (1.3) with initial data $\xi \in C$, $0 \le q \le \varepsilon$. Suppose that*

$$\Phi_\varepsilon(x, y) = \sum_{i=1}^n \sum_{l=1}^L \alpha_{il}(y_i^{\beta_l} - \eta_i^{-1}e^{b\varepsilon}x_i^{\beta_l}) \quad (x, y \in \mathbb{R}^n) \tag{2.9}$$

where α_{il} and β_l $(1 \le i \le n, 1 \le l \le L)$ are nonnegative constants; b is given by (2.1) and η_i $(1 \le i \le n)$ is given by (2.2). Then

$$\int_0^t e^{qs}\Phi_\varepsilon(x(s), y(s))\,ds \le const \quad \text{for } t \ge 0.$$

In this paper, "const" always denotes a positive constant with different values in different places.

For the convenience of readers, let us cite some useful results on M-matrices [21].

Definition 2.2. If $Q = [q_{ij}] \in \mathbb{R}^{n\times n}$ satisfies $q_{ij} \le 0 < q_{ii}$ $(i \ne j,\ i, j = 1, 2, \ldots, n)$, and all eigenvalues of Q have positive real parts, then Q is called an M-matrix.

Denote that $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n : x_i > 0 (1 \le i \le n)\}$. We assume that $x \gg 0 \Leftrightarrow x \in \mathbb{R}^n_{++}$.

Lemma 2.3. *Let $Q = [q_{ij}] \in \mathbb{R}^{n\times n}$ satisfy $q_{ij} \le 0 < q_{ii}$ $(i \ne j,\ i, j = 1, 2, \ldots, n)$. Then the following statements are equivalent:*

(i) *Q is an M-matrix.*
(ii) *There exists $c \in \mathbb{R}^n_{++}$ such that $Qc \gg 0$.*
(iii) *All of the leading principal minors of Q are positive.*

Remark 2.4. Lemma 2.3 implies that: (i) if Q is an M-matrix, then both Q^T and $\alpha Q(\alpha > 0)$ are M-matrices; (ii) M-matrix Q remains as an M-matrix under tiny disturbances if the condition $q_{ij} \le 0 < q_{ii}$ $(j \ne i,\ i, j = 1, 2, \ldots, n)$ holds. (iii) for any given $c \in \mathbb{R}^n_{++}$, $\mathrm{diag}(c)$ is an M-matrix.

We assume that coefficients of Eq. (1.3) satisfy the following linear growth conditions:

$$|g_i(y)| \le \rho_i|y_i|; \tag{2.10}$$

$$|\sigma_i(t, x, y, u, v)|^2 \le \sum_{j=1}^n (\kappa_{ij}x_j^2 + \tau_{ij}y_j^2 + \lambda_{ij}u_j^2 + \mu_{ij}v_j^2), \tag{2.11}$$

where $t \ge 0$; $x, y, u, v \in \mathbb{R}^n$; $\rho_i, \kappa_{ij}, \tau_{ij}, \lambda_{ij}$ and μ_{ij} are nonnegative constants with $i, j = 1, 2, \ldots, n$; $\sigma^T = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ with $\sigma_i(1 \le i \le n)$ being R^m−valued functions.