第 23 卷 第 5 期 电子与信息学报 Vol.23 No.5
2001 年 5 月 JOURNAL OF ELECTRONICS AND INFORMATION TECHNOLOGY May 2001

随机 Hopfield 神经网络的定量分析 [1]

金 聪

(湖北大学数学与计算机科学学院 武汉 430062)

摘 要 该文探讨了实际使用 Hopfield 神经网络（HNN）时噪声的影响. 由于噪声的客观存在, 我们首先证明了随机 Hopfield 神经网络（SHNN）轨道的期望关于时间是一致有界的. 之后, 为了实际设计神经网络的需要, 我们对含有噪声的 HNN 和与其对应的一般 HNN 之间随机输入误差的估计进行了研究. 利用所得的结论, 我们可以对设计空间进行控制, 使得所设计的网络满足我们希望获得的各种性能要求.

关键词 随机 Hopfield 网络, 噪声, 随机输入误差, 一致有界

中图号 TN-052

1 引 言

最近几年来, Hopfield 神经网络 [1] 的理论研究取得了飞速的发展, 获得了很多有意义的成果. 这些成果已在模式识别、自动控制等诸多领域取得了令人瞩目的进展. 然而, 由于输入噪声的存在, 使得在使用 HNN 的过程中, 难以对噪声的存在进行有效的控制. 文献 [2] 研究了神经网络动力系统在白噪声小扰动下的随机非线性效应, 重点研究了推广的 Hopfield 模型在权值扰动下的网络结构稳定性问题. 本文的研究内容不同于文献 [2]. 本文将对含有输入白噪声的随机神经网络进行研究, 尤其对含噪声的随机神经网络与不含随机扰动的神经网络之间的输入误差进行研究. 所得结果可以用来设计 HNN, 使其满足设计者的各种特殊性能要求.

2 模型描述

随机 Hopfield 神经网络可由如下微分方程描述:

$$\frac{\mathrm{d}V_i(t)}{\mathrm{d}t} = -a_iV_i(t) + \sum_{j=1}^{n} w_{ij}f(V_i(t)) + I_i(t) + \xi_i(t), \quad 1 \le i \le n \tag{1}$$

而微分方程

$$\frac{\mathrm{d}\tilde{V}_i(t)}{\mathrm{d}t} = -a_i\tilde{V}_i(t) + \sum_{j=1}^{n} w_{ij}f(\tilde{V}_i(t)) + I_i(t), \quad 1 \le i \le n \tag{2}$$

称为与 (1) 式对应的 Hopfield 神经网络. 其中 $V_i(t)$ 是第 i 个神经元的状态; w_{ij} 是与时间无关的第 j 个神经元到第 i 个神经元的连接权, w_{ij} 不一定对称; a_i 是恒为正的有界时间常数; $I_i(t)$ 是第 i 个节点的时变输入; $f(\cdot)$ 是 Sigmoid 型函数, 本文取作 $f(x) = 1/(1+e^{-x})$; 而 $\xi_i(t)$ 是均值为 0 的高斯白噪声.

注意到 $V_i(t)$ 是一个随机变量, 其不确定的程度是由附加的噪声 $\xi_i(t)$ 引起的. 从严格的数学观点出发, (1) 式可以看成是 Itô 随机微分方程 [3]

[1] 1999-05-09 收到, 1999-10-20 定稿

$$\mathrm{d}V_i(t) = [-a_iV_i(t) + \sum_{j=1}^{n} w_{ij}f(V_i(t)) + I_i(t)]\mathrm{d}t + \mathrm{d}\beta_i(t) \tag{3}$$

其中 $\beta_i(t)(1 \le i \le n)$ 是第 i 个神经元的零均值 Wiener 过程。

3 主要结果

引理 1 对于微分方程 $\dot{x}_i(t) = -ax_i(t) + g(x(t))$ ， $1 \le i \le n$ ，其中 $a > 0, x(t) \in R^n$ 且 $-\infty < M_1 \le g(\cdot) \le M_2 < \infty$ ，则有

$$\frac{M_1}{a} \le x_i(t) \le \frac{M_2}{a}$$

证明请见文献 [4] 。

定理 1 若 $I_i(t)$ 是关于时间 t 的一致有界函数，且 $\sum_{j=1}^{n}|w_{ij}| = a < \infty$ ，则系统 (3) 式对任意神经元 $V_i(t)$ ，其数学期望 $E(V_i(t))$ 关于 t 是一致有界的。

证明 为了获得 $V_i(t)$ 的数学期望，对 (3) 式两边积分得

$$V_i(t) = V_i(0) + \int_0^t [-a_iV_i(\tau) + \sum_{j=1}^{n} w_{ij}f(V_i(\tau)) + I_i(\tau)]\mathrm{d}\tau + \int_0^t \mathrm{d}\beta_i(t)$$

两边取数学期望，并注意到 Fubini 定理得

$$E(V_i(t)) = E(V_i(0)) + \int_0^t [-a_iE(V_i(\tau)) + \sum_{j=1}^{n} w_{ij}E(f(V_i(\tau))) + I_i(\tau)]\mathrm{d}\tau$$

两边求关于时间 t 的导数，得

$$\dot{E}(V_i(t)) = -a_iE(V_i(t)) + \sum_{j=1}^{n} w_{ij}E(f(V_i(t))) + I_i(t) \tag{4}$$

由 $f(\cdot)$ 的定义及 $I_i(t)$ 关于 t 的一致有界性知 $\sum_{j=1}^{n} w_{ij}E(f(V_i(t))) + I_i(t)$ 是关于时间 t 一致有界的，所以由引理 1 知 $E(V_i(t))$ 关于时间一致有界。 证毕

对于与 (1) 式对应的 Hopfield 神经网络 (2) 式

$$\frac{\mathrm{d}\tilde{V}_i(t)}{\mathrm{d}t} = -a_i\tilde{V}_i(t) + \sum_{j=1}^{n} w_{ij}f(\tilde{V}_i(t)) + I_i(t), \quad 1 \le i \le n \tag{5}$$

以下我们要研究 $V_i(t)$ 与 $\tilde{V}_i(t)$ 之间误差的定量关系， $1 \le i \le n, t \ge 0$ 。为此，先给出几个定义。

定义 1 若对于 Hopfield 神经网络 (2) 式，存在常数矩阵 P 使得

$$\tilde{Y}(t) = P\tilde{V}(t)$$

则称 $\tilde{Y}(t)$ 是 (2) 式的输出响应， P 称为输出矩阵。

定义 2 若对于随机 Hopfield 神经网络 (1) 式，存在常数矩阵 P 使得

$$Y(t) = PV(t)$$

则称 $Y(t)$ 是系统 (1) 式的输出响应。

定义 3　对于系统 (1) 和 (2) 式，称

$$d_i(t) = E(V_i(t)) - \tilde{V}_i(t), \quad 1 \le i \le n \tag{6}$$

为系统 (1) 式在 t 时刻的随机输入误差，由此易知系统 (1) 式在 t 时刻的均方误差为

$$E(V_i(t) - \tilde{V}_i(t))^2 = \sigma_i^2(t) + d_i^2(t)$$

这里 $\sigma_i^2(t)$ 是 $V_i(t)$ 的方差。

记 $\lambda_{\max}(\cdot)$ 表示矩阵 $(\cdot)$ 的最大特征值。

定理 2　若系统 (1) 式满足如下条件:

(1) $I_i(t)$ 是关于时间 t 的一致有界函数;

(2) $\sum_{j=1}^{n} |w_{ij}| = a < \infty$;

(3) $\lambda_{\max}(B(t)W + W^T B(t) - 2A) < 0, \ t \ge 0$; 其中 $B(t) = \mathrm{diag}(b_1(t), \cdots, b_n(t))$ ，而

$$0 < b_i(t) = \frac{f(E(V_i(t))) - f(\tilde{V}_i(t))}{d_i(t)} \le \frac{1}{4}, \qquad 1 \le i \le n$$

则必有 $\lim_{t\to\infty} \|d(t)\| \le ak_1\sqrt{n}/k_2$ 。这里正常数 k_1 和 k_2 是与系统 (1) 式相对应的状态转移矩阵的衰减速率。

证明 由 (6) 式可知 $\tilde{V}_i(t) = E(V_i(t)) - d_i(t)$ ，则由 (5) 式我们可以获得

$$\frac{\mathrm{d}}{\mathrm{d}t}(E(V_i(t)) - d_i(t)) = -a_i(E(V_i(t)) - d_i(t)) + \sum_{j=1}^{n} w_{ij} f(E(V_i(t)) - d_i(t)) + I_i(t) \tag{7}$$

将 $f(V_i(t))$ 在 $E(V_i(t))$ 处展开成 Toylor 级数，有 $f(V_i(t)) = f(E(V_i(t))) + \gamma_i(t)$ ，其中 $\gamma_i(t)$ 是余项，由此获得

$$E(f(V_i(t))) = f(E(V_i(t))) + E(\gamma_i(t)) \tag{8}$$

因为 $0 < f(\cdot) < 1$ ，所以 $-1 < \gamma_i(t) < 1$ ，易知 (4) 式成立，故将 (8) 式代入 (4) 式得

$$\begin{aligned}\dot{E}(V_i(t)) &= -a_i E(V_i(t)) + \sum_{j=1}^{n} w_{ij}[f(E(V_i(t))) + E(\gamma_i(t))] + I_i(t) \\ &= -a_i E(V_i(t)) + \sum_{j=1}^{n} w_{ij} f(E(V_i(t))) + \sum_{j=1}^{n} w_{ij} E(\gamma_i(t)) + I_i(t) \\ &= -a_i E(V_i(t)) + \sum_{j=1}^{n} w_{ij} f(E(V_i(t))) + q_i(t) + I_i(t) \end{aligned} \tag{9}$$

其中 $q_i(t) = \sum\limits_{j=1}^{n} w_{ij} E(\gamma_i(t))$ 。将 (9) 式代入 (7) 式可得

$$\begin{aligned}\frac{\mathrm{d}d_i(t)}{\mathrm{d}t} &= \dot{E}(V_i(t)) - \frac{\mathrm{d}}{\mathrm{d}t}[E(V_i(t)) - d_i(t)] \\ &= -a_i d_i(t) + \sum_{j=1}^{n} w_{ij}[f(E(V_i(t))) - f(E(V_i(t)) - d_i(t))] + q_i(t) \\ &= -a_i d_i(t) + \sum_{j=1}^{n} w_{ij}[f(E(V_i(t))) - f(\tilde{V}_i(t))] + q_i(t) \end{aligned} \tag{10}$$

注意到 $f(\cdot)$ 的单调递增及有界性，由中值定理可得

$$f(E(V_i(t))) - f(\tilde{V}_i(t)) = b_i(t)(E(V_i(t)) - \tilde{V}_i(t)) = b_i(t)d_i(t)$$

由于 $0 < f'(\cdot) \le 1/4$，所以 $0 < b_i(t) = \frac{f(E(V_i(t))) - f(\tilde{V}_i(t))}{d_i(t)} \le \frac{1}{4}$， $1 \le i \le n$。将上式代入 (10) 式得 $\frac{\mathrm{d}d_i(t)}{\mathrm{d}t} = -a_i d_i(t) + b_i(t)\sum_{j=1}^{n} w_{ij} d_i(t) + q_i(t)$， $1 \le i \le n$， $t \ge 0$，写成矩阵形式即

$$\dot{d}(t) = (-A + B(t)W)d(t) + q(t) \tag{11}$$

其中 $d(t) = [d_1(t), \cdots, d_n(t)]^T$， $A = \mathrm{diag}[a_1, \cdots, a_n]$， $B(t) = \mathrm{diag}[b_1(t), \cdots, b_n(t)]$， $q(t) = [q_1(t), q_2(t), \cdots, q_n(t)]^T$ 。考虑齐次方程：

$$\dot{\tilde{d}}(t) = [-A + B(t)W]\tilde{d}(t) \tag{12}$$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\|P\tilde{d}(t)\| &= \frac{\mathrm{d}}{\mathrm{d}t}(\tilde{d}^T(t)P^T P\tilde{d}(t)) = \dot{\tilde{d}}^T(t)P^T P\tilde{d}(t) + \tilde{d}^T(t)P^T P\dot{\tilde{d}}(t) \\ &= \tilde{d}^T(t)[-A + B(t)W]^T P^T P\tilde{d}(t) + \tilde{d}^T(t)P^T P[-A + B(t)W]\tilde{d}(t) \\ &= \tilde{d}^T(t)[(-A + B(t)W)^T P^T P + P^T P(-A + B(t)W)]\tilde{d}(t) \\ &= \tilde{d}^T(t)C(t)\tilde{d}(t)\end{aligned}$$

注意到 $C(t)$ 是对称矩阵，由文献 [5] 中的 Rayleigh-Ritz 不等式有

$$\frac{\mathrm{d}}{\mathrm{d}t}\|P\tilde{d}(t)\|^2 \le \lambda_{\max}(C(t))\|\tilde{d}(t)\|^2$$

设对于非零输入误差有 $\|P\tilde{d}(t)\|^2 > 0$，从而得到

$$\frac{1}{\|P\tilde{d}(t)\|^2} \cdot \frac{\mathrm{d}}{\mathrm{d}t}\|P\tilde{d}(t)\|^2 \le \lambda_{\max}(C(t)) \cdot \frac{\|\tilde{d}(t)\|^2}{\|P\tilde{d}(t)\|^2}$$

记 $\frac{\|\tilde{d}(t)\|^2}{\|P\tilde{d}(t)\|^2} = k(P,t)$，对上式两边取积分得

$$\ln\|P\tilde{d}(t)\|^2 - \ln\|Pd(0)\|^2 \le \int_0^t \lambda_{\max}(C(\tau))k(P,\tau)\mathrm{d}\tau$$

于是 $\ln\|P\tilde{d}(t)\| \le \ln\|Pd(0)\| + \frac{1}{2}\int_0^t \lambda_{\max}(C(\tau))k(P,\tau)\mathrm{d}\tau$，从而获知

$$\|P\tilde{d}(t)\| \leq \|Pd(0)\| \cdot e^{\frac{1}{2}\int_0^t \lambda_{\max}(C(\tau))k(P,\tau)\mathrm{d}\tau}, \quad t \geq 0$$

由文献 [6] 知系统 (11) 式的状态方程解的表达式为

$$d(t) = \Phi(t,0)d(0) + \int_0^t \Phi(t,\tau)q(\tau)\mathrm{d}\tau, \quad t \geq 0$$

这里 $\Phi(\cdot,\cdot)$ 是该系统的状态转移矩阵，因此

$$\begin{aligned}\|d(t)\| &= \left\|\Phi(t,0)d(0) + \int_0^t \Phi(t,\tau)q(\tau)\mathrm{d}\tau\right\| \leq \|\Phi(t,0)d(0)\| + \left\|\int_0^t \Phi(t,\tau)q(\tau)\mathrm{d}\tau\right\| \\ &\leq \|\Phi(t,0)\| \cdot \|d(0)\| + \int_0^t \|\Phi(t,\tau)\| \cdot \|q(\tau)\|\mathrm{d}\tau\end{aligned}$$

因为 $\|q(t)\| = \|\sum_{j=1}^n w_{ij}E(\gamma_i(t))\| \leq \sum_{j=1}^n |w_{ij}| \cdot \|E(\gamma_i(t)\| \leq \sum_{j=1}^n |w_{ij}| \cdot \sqrt{n} = a\sqrt{n}$ ，所以

$$\|d(t)\| \leq \|\Phi(t,0)\| \cdot \|d(0)\| + a\sqrt{n} \cdot \int_0^t \|\Phi(t,\tau)\|\mathrm{d}\tau \tag{13}$$

由文献 [6] 及定理 2 中条件 (3) 推知系统 (12) 式是一致指数稳定的，于是存在与 τ 无关的正常数 k_1 和 k_2 ，使得

$$\|\Phi(t,\tau)\| \leq k_1 e^{-k_2(t-\tau)}, \quad t \geq \tau$$

将此不等式代入 (13) 式得

$$\begin{aligned}\|d(t)\| &\leq k_1 e^{-k_2 t} \cdot \|d(0)\| + a\sqrt{n} \cdot \int_0^t k_1 e^{-k_2(t-\tau)}\mathrm{d}\tau \\ &= k_1 e^{-k_2 t} \cdot \|d(0)\| + a\sqrt{n} \cdot \frac{k_1}{k_2}(1 - e^{-k_2\tau}) \\ &= a\sqrt{n} \cdot k_1/k_2 + k_1(\|d(0)\| - a\sqrt{n}/k_2)e^{-k_2 t}\end{aligned}$$

两边取极限得知 $\lim_{t\to\infty} \|d(t)\| \leq ak_1\sqrt{n}/k_2$ 。　　证毕

4 结　论

在这篇文章中，我们讨论了输入噪声为白噪声情况下随机 Hopfield 神经网络的行为特征。由分析可知，当时间趋于无穷时，SHNN 和与其对应的 HNN 之间的随机输入误差的测度并不随时间的变化而变大，而是与 $\sqrt{n}$ 成正比。这表明，随着神经网络规模的扩大，误差的影响也在扩大。这与人们一般的认识是一致的。本文给出了定量分析结果，这在实际设计 Hopfield 神经网络中有重要的指导意义。

参　考　文　献

[1]　徐秉铮，张百灵，韦岗，神经网络理论与应用，广州，华南理工大学出版社， 1994, 第二章.

[2]　廖桂生，焦李成，保铮，神经网络随机扰动稳定性研究，电子科学学刊， 1992, 14(4), 416–419.

[3] 武宝亭，李庆士，杨跃武，随机过程与随机微分方程，成都，电子科技大学出版社，1994, 第八章.

[4] M. A. Colen, S. Grossberg, Absolute stability of global pattern formation and parallel memory storage by competitive neural networks, IEEE Trans. on Syst., Man, Cybern., 1983, SMC-13(5), 815-826.

[5] 王松桂，贾忠贞，矩阵论中不等式，合肥，安徽教育出版社，1994, 第四章.

[6] 段广仁，线性系统理论，哈尔滨，哈尔滨工业大学出版社，1996, 第三章，第五章.

THE QUANTITATIVE ANALYSIS OF STOCHASTIC HOPFIELD NEURAL NETWORK MODEL

Jin Cong

(*College of Mathematics and Computer Science, Hubei University, Wuhan* 430062, China)

Abstract In this paper, the effect of input noise on the typical stochastic Hopfield neural network model is discussed. It is shown that the expectation of the stochastic HNN of the trajectory is uniformly bounded over time. For practical design purposes, the stochastic input error estimates for the stochastic HNN with respect to the corresponding deterministic HNN is derived. In addition, the designer can use these results to constrain the design space so that the achieved design satisfies the performance specifications whenever possible.

Key words Stochastic Hopfield neural network, Noise, Stochastic input error, Uniform bounded

金 聪： 女，1960 年生，理学硕士，副教授，主要研究方向为神经网络、智能信息处理及并行计算，已在国内外发表学术论文 20 多篇.