第14卷 第4期　　电子科学学刊　　Vol. 14, No. 4
1992年7月　　JOURNAL OF ELECTRONICS　　July, 1992

神经网络随机扰动稳定性研究*

廖桂生　焦李成　保铮
(西安电子科技大学电子工程所，西安 710071)

摘要　本文基于随机微分方程理论，严格地分析了一类广义的 Hopfield 连续时间神经网络在白噪声扰动下的稳定性，并建立了相应的稳定性判据和网络的设计准则。

关键词　Hopfield 神经网络；随机微分方程；稳定性；Wiener 过程。

1. 引言

目前应用神经网络，无论是联想存贮记忆，还是神经优化计算，都毫无例外地利用了网络系统的稳定吸引子的性质。对 Hopfield[1] 神经网络的稳定性研究，目前已有许多工作，讨论已由简单的零对角对称连接矩阵、异步方式的稳定性发展到任意的连接矩阵、同步方式的稳定性。所有这些讨论，大多限于状态空间初始扰动的稳定性，而对神经网络本身的随机扰动稳定性，或者网络结构的稳定性，尽管在实际实现神经网络中很重要，还较少涉及。

本文研究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 A. N. Michel 等人[2,3]所推广的 Hopfield 模型在权值扰动下的网络结构稳定性问题。

2. 神经网络模型

为了便于叙述，首先回顾 Hopfield 模型，然后给出本文的模型。文献上的常用符号不另作说明。

(1) Hopfield 神经网络模型　　Hopfield 连续时间神经网络模型如图1所示，用微分方程系统表示则为

$$C_i\dot{u}_i = \sum_{j=1}^{N} T_{ij}v_j - u_i/\tau_i + I_i(t), \qquad i = 1,\cdots,N \tag{1}$$

其中 $C_i > 0$; $T_{ij} = 1/R_{ij}$, $R_{ij} \in R = (-\infty,\infty)$; $1/\tau_i = 1/R_i + \sum_{j=1}^{N}|T_{ij}|$, $R_i > 0$; I_i: $R^+ = [0,\infty) \to R$, I_i 是连续的; $\dot{u}_i = du_i/dt$; $v_i = g_i(u_i)$, g_i: $R \to (-1,1)$, g_i 是连续可微且严格单增的，且 $g_i(0) = 0$; 各参数的物理意义见文献[1,2]。令

$$X = (x_1,\cdots,x_N)^T,\ x_i = u_i,\ i = 1,\cdots,N$$

$$B = \mathrm{diag}(b_1,\cdots,b_N),\ b_i = 1/(\tau_i C_i),\ i = 1,\cdots,N$$

1991.04.01 收到，1992.03.13 定稿。
* 国家自然科学基金资助课题。

$$A = [A_{ij}]_{N\times N},\ A_{ij} = T_{ij}/C_i,\ i,j = 1,\cdots,N$$
$$G = [g_1(x_1),\cdots,g_N(x_N)]^T,\ U(t) = (I_1(t),\cdots,I_N(t))^T$$

则(1)式等价于(2)式

$$\dot{X} = -BX + AG(X) + U(t) \tag{2}$$

(2) 随机扰动的 Hopfield 神经网络　(2) 式中参数 A, B 在实际电路中可能存在小的随机扰动，即

$$\acute{A} = A + \Delta A \tag{3}$$
$$\acute{B} = B + \Delta B \tag{4}$$

将(2)式中 A,B 分别用 $\acute{A},\acute{B}$ 代替，则有

$$\dot{X} = -BX + AG(X) + U(t) + \varepsilon(X) \tag{5}$$

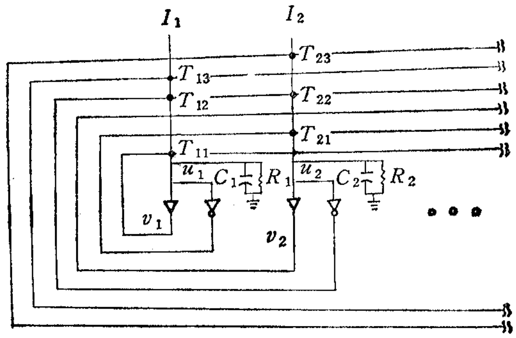

图 1　Hopfield 神经网络模型

其中

$$\varepsilon(X) = -\Delta BX + \Delta AG(X) \tag{6}$$

一般地，$\varepsilon(X)$ 是一随机过程，本文设

$$\varepsilon(X) = \sigma(X)\xi_t \tag{7}$$

其中 $\sigma(X)$: $R^N \to R^{N\times N}$ 为扩散矩阵，$\xi_t \in R^N, \{\xi_t^i, t\in T\}$ 是一个具有独立增量的规范的 Wiener 过程.

由上述讨论得知，对应于(2)式的随机扰动下的 Hopfield 神经网络模型为

$$\dot{X} = -BX + AG(X) + U(t) + \sigma(X)\xi_t \tag{8}$$

扩散矩阵 $\sigma(X)$ 取决于扰动量和状态变量与平衡点之间的距离. 不失一般性，我们只需研究(2)式的平凡平衡点情形. 因此，提出本文的假设

假设 1　存在正数 $\Gamma > 0$ 和N阶对角矩阵 $\Phi = \mathrm{diag}(\varphi_1,\cdots,\varphi_N)$，其中 $\varphi_i > 0$，$i = 1,\cdots,N$，使得

$$\|\Phi\sigma(X)\| = \{\mathrm{tr}[\sigma(X)^T\Phi^2\sigma(X)]\}^{1/2} \leqslant \Gamma\|X\|_2 \tag{9}$$

其中 $\|\cdot\|_2$ 是欧几里德范数.

3. 稳定性分析

Michel 等人对神经网络模型(2)式的稳定性进行了深入的研究，得到了较一般性的稳定性结果. 本文在他们的基础上进一步研究了对应的随机神经网络(8)式的稳定性. 与文献[2]一样，只需研究网络的平凡平衡点的稳定性. 所以，以后假定 $U(t) = 0$.

定理 1[2]　神经网络系统(2)式的平衡点 $X = 0$ 是指数稳定的，如果如下条件均成立：

条件 1　存在正常数 $r_i > 0$，使得对所有的 $|x_i| < r_i$，$|x_j| < r_j$，$i,j = 1, 2,\cdots,N$，有

$$x_iA_{ij}G(x_j) \leqslant |x_i|a_{ij}|x_j| \tag{10}$$

其中 a_{ij}，$i,j = 1,\cdots,N$ 为常数.

条件 2　存在N阶对角矩阵 $\alpha = \mathrm{diag}(\alpha_1,\cdots,\alpha_N)$，$\alpha_i > 0$，$i = 1,\cdots,N$，使得

检验矩阵 $\boldsymbol{S}=[S_{ij}]_{N\times N}$:

$$S_{ij}=\begin{cases}\alpha_i(-b_i+a_{ii}), & i=j\\(\alpha_i a_{ij}+\alpha_j a_{ji})/2, & i\neq j\end{cases} \tag{11}$$

是负定的. 其中 b_i 由(2)式给出, a_{ij} 由条件 1 给定.

定理 2 神经网络系统(8)式的平衡点 $\boldsymbol{X}=0$ 是按概率 1 均方指数稳定的, 如果定理 1 中两个条件关于系统(8)式成立,而且存在正数 Γ 使扩散矩阵 $\boldsymbol{\sigma}(\boldsymbol{X})$ 满足

(1) $$\Gamma<-\lambda_M(\boldsymbol{S}) \tag{12}$$

(2) $$(1/2)\mathrm{tr}[\boldsymbol{\sigma}(\boldsymbol{X})^T\boldsymbol{\Phi}\boldsymbol{\sigma}(\boldsymbol{X})]\leqslant\Gamma\|\boldsymbol{X}\|_2^2 \tag{13}$$

其中 $\lambda_M(\boldsymbol{S})$ 表示实对称矩阵 $\boldsymbol{S}$ 的最大特征值,对角矩阵 $\boldsymbol{\Phi}=\mathrm{diag}(\alpha_1,\cdots,\alpha_N)$, 这里的 $\boldsymbol{S}$ 和 α_i, $i=1,\cdots,N$ 由定理 1 中条件 2 给出.

证明 选取 Lyapunov 函数

$$V(\boldsymbol{X})=\frac{1}{2}\sum_{i=1}^{N}\alpha_i x_i^2=\frac{1}{2}\boldsymbol{X}^T\boldsymbol{\Phi}\boldsymbol{X} \tag{14}$$

由条件 2 知 $\alpha_i>0$, $i=1,\cdots,N$, 显然 $V(\boldsymbol{X})$ 是正定的. $V(\boldsymbol{X})$ 沿随机微分方程(8)式的解关于时间 t 的微分定义如下

$$DV(\boldsymbol{X})=\lim_{\delta\to0^+}E_{x,t}[V(\boldsymbol{X}_{t+\delta})-V(\boldsymbol{X})]/\delta \tag{15}$$

由于假设 $\boldsymbol{\xi}_t$ 是具有独立增量的规范的 Wiener 过程,故

$$\begin{aligned}DV(\boldsymbol{X})=&\nabla_t V(\boldsymbol{X})+[\nabla_x V(\boldsymbol{X})]^T[-\boldsymbol{BX}+\boldsymbol{AG}(\boldsymbol{X})]\\&+(1/2)\mathrm{tr}[\boldsymbol{\sigma}(\boldsymbol{X})^T\nabla_{xx}V(\boldsymbol{X})\boldsymbol{\sigma}(\boldsymbol{X})]\end{aligned} \tag{16}$$

将(14)式代入(16)式,得

$$DV(\boldsymbol{X})=-\boldsymbol{X}^T\boldsymbol{\Phi}\boldsymbol{BX}+\boldsymbol{X}^T\boldsymbol{\Phi}\boldsymbol{AG}(\boldsymbol{X})+(1/2)\mathrm{tr}[\boldsymbol{\sigma}(\boldsymbol{X})^T\boldsymbol{\Phi}\boldsymbol{\sigma}(\boldsymbol{X})] \tag{17}$$

$$=-\sum_{i=1}^{N}\alpha_i b_i x_i^2+\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i x_i A_{ij}G_j(x_j)+\frac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}(\boldsymbol{X})^T\boldsymbol{\Phi}\boldsymbol{\sigma}(\boldsymbol{X})] \tag{18}$$

由条件 1, 即(10)式,有

$$DV(\boldsymbol{X})\leqslant\sum_{i=1}^{N}\alpha_i\left(-b_i x_i^2+|x_i|\sum_{j=1}^{N}a_{ij}|x_j|\right)+\frac{1}{2}\mathrm{tr}[\boldsymbol{\sigma}(\boldsymbol{X})^T\phi\boldsymbol{\sigma}(\boldsymbol{X})] \tag{19}$$

令 $r=\min\limits_i(r_i)$, $\boldsymbol{R}=[r_{ij}]_{N\times N}$, 其中

$$r_{ij}=\begin{cases}\alpha_i(-b_i+a_{ii}), & i=j\\\alpha_i a_{ij}, & i\neq j\end{cases}$$

则当 $\|\boldsymbol{X}\|_2<r$ 时,有

$$DV(\boldsymbol{X})\leqslant\boldsymbol{X}^T\boldsymbol{RX}+(1/2)\mathrm{tr}[\boldsymbol{\sigma}(\boldsymbol{X})^T\boldsymbol{\Phi}\boldsymbol{\sigma}(\boldsymbol{X})] \tag{20}$$

注意到

$$\boldsymbol{X}^T\boldsymbol{RX}=\boldsymbol{X}^T[(\boldsymbol{R}+\boldsymbol{R}^T)/2]\boldsymbol{X}=\boldsymbol{X}^T\boldsymbol{SX}\leqslant\lambda_M(\boldsymbol{S})\|\boldsymbol{X}\|_2^2 \tag{21}$$

把(13)式和(21)式代入(20)式,得

$$DV(\boldsymbol{X})\leqslant\lambda_M(\boldsymbol{S})\|\boldsymbol{X}\|_2^2+\Gamma\|\boldsymbol{X}\|_2^2=(\lambda_M(\boldsymbol{S})+\Gamma)\|\boldsymbol{X}\|_2^2 \tag{22}$$

因为 $\boldsymbol{S}$ 是负定的,故 $\lambda_M(\boldsymbol{S})<0$. 由(14)式、(12)式和(22)式推出在原点 $\boldsymbol{X}=0$ 的某邻域内,有

$$C_1\|\boldsymbol{X}\|_2^2\leqslant V(\boldsymbol{X})\leqslant C_2\|\boldsymbol{X}\|_2^2,\ DV(\boldsymbol{X})\leqslant-C_3\|\boldsymbol{X}\|_2^2 \tag{23}$$

其中$C_1 = (1/2)\min_i(\alpha_i) > 0$，$c_2 = (1/2)\max_i > 0$，$c_3 = -(\lambda_M(\boldsymbol{S}) + \Gamma) > 0$，因此，根据随机微分方程理论[4]，平衡点 $\boldsymbol{X} = 0$ 是按概率 1 均方指数稳定的。　　证毕

定理 2 告诉我们，确定型理想的神经网络系统(2)式的稳定裕度越大，即要求检验矩阵 $\boldsymbol{S}$ 的最大特征值越负，则在随机扰动下越能保证网络稳定；同样，系统(2)式的扰动幅度越小，即扩散矩阵 $\boldsymbol{\sigma}(\boldsymbol{X})$ 的"加权"范数越小，则网络越稳定。这两种情形与我们对神经网络已有的直观理解是吻合的。定理 2 为我们提供了设计网络的新的可靠依据，由此可设计出具有相当稳定性的神经网络。

为了保证定理 2 成立，要求扰动有一定的动态范围，即扩散矩阵 $\boldsymbol{\sigma}(\boldsymbol{X})$ 满足假设 1. 从定理 2 的证明过程可清楚地看到，一般地，当网络的状态距离平衡点越远时，网络越不稳定，对随机扰动的幅度要求就越宽；相反，当网络的状态趋于稳定平衡点时，网络的随机扰动越小。实际情况亦如此，当网络经动力学演化达到稳定平衡点时，热噪声扰动趋于足够小，以至可以忽略。由此也可以看出，我们对扩散矩阵所作的假设 1 是合理的。

4. 结论

本文从随机微分方程角度出发，严格地分析了广义的 Hopfield 连续神经网络在噪声扰动下的稳定性，并建立了相应的稳定性判据和网络的设计准则，这为用 VLSI 实现人工神经网络提供了新的理论基础和设计准则。本文建立的分析方法也可以用于研究一般的神经网络随机扰动稳定性问题。

参　考　文　献

[1] J. J. Hopfield, *Proc. Natl. Acad. Sci.* USA, **81** (1984) 3, 3088—3092.
[2] A. N. Michel, J.A. Farrell, W. Porod, *IEEE Trans, on CAS*, **CAS-36** (1989)2, 229—243.
[3] J. Li, A. N. Michel, W. Porod, *IEEE Trans. on CAS*, **CAS-36**(1989)5, 713—731.
[4] 胡宣达，随机微分方程稳定性理论，南京大学出版社，南京，1986 年，第 63—135页.

ON STABILITY STUDY OF NEURAL NETWORKS WITH RANDOM PERTURBATION

Liao Guisheng　　Jiao Lichen　　Bao Zheng
(Xidian University, Xi'an 710071)

Abstract　　Based on the theory of stochastic differential equation, the stability of a kind of generalized continuous-time Hopfield neural network with white noise perturbation is studied, and the related stability criteria and design requirements of neural networks are established.

Key words　　Hopfield neural network; Stochastic differential equation; Stability; Wiener process