

离散 Hopfield 神经网络的稳定性研究

马润年¹, 张 强¹, 许 进²

(1. 西安电子科技大学电子工程研究所, 陕西西安 710071; 2. 华中科技大学系统科学研究所, 湖北武汉 430074)

摘 要: 离散 Hopfield 神经网络的稳定性是网络应用的基础. 文中主要研究非对称离散 Hopfield 神经网络的异步、同步、部分同步演化方式的稳定性, 并给出了一些新的稳定性条件, 所获结果推广了一些已有的结论.

关键词: 离散 Hopfield 神经网络; 稳定性; 能量函数

中图分类号: TP183 **文献标识码:** A **文章编号:** 0372-2112 (2002) 07-1089-03

Stability Study of Discrete Hopfield Neural Networks

MA Run-nian¹, ZHANG Qiang¹, XU Jin²

(1. Electronic Engineering Research Institute, Xidian University, Xi'an, Shaanxi 710071, China;

2. System Science Research Institute, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China)

Abstract: The stability of Hopfield neural network is the foundation of the network's applications. The stability of asymmetric discrete Hopfield neural network is mainly studied and some new simple stability conditions are given in serial, parallel and partially parallel updating mode. The obtained results generalize some existing results.

Key words: discrete Hopfield neural networks; stability; energy function

1 引言

离散 Hopfield 神经网络(以下简称 DHNN)在联想记忆、模式识别和组合优化等方面有着广泛应用的一类著名的网络^[1~6]. 在这些应用中都需要网络是稳定的, 即网络的任何轨迹最后都收敛到稳定点(或不动点), 因此 DHNN 的稳定性是网络各种应用的基础, 从而受到许多学者的极大关注. 本文主要在文[1~9]的基础上进一步研究非对称且阈值不一定为零的 DHNN 在各种演化方式下的稳定性, 并给出了几个简单的稳定性结论, 所获结果推广了部分原有的稳定性结论.

2 基本概念

具有 n 个神经元的离散 Hopfield 神经网络, 其拓扑结构可以由一个 $n \times n$ 阶矩阵 $W = (w_{ij})_{n \times n}$ 和一个 n 维列向量 $\theta = (\theta_1, \dots, \theta_n)^T$ 所唯一确定, 并记为 $N = (W, \theta)$. 若用 $x_i(t)$ 表示神经元 i 在时刻 t 所处的状态, 并且只有两种: 兴奋用 $x_i(t) = 1$ 表示和抑制用 $x_i(t) = -1$ 表示, $t = 0, 1, 2, \dots, j$. DHNN 的演化方程为:

$$x_i(t+1) = \text{sgn}\{H_i(t)\} = \begin{cases} 1, & H_i(t) \geq 0 \\ -1, & H_i(t) < 0 \end{cases} \quad (1)$$

其中 $H_i(t) = \sum_{j=1}^n w_{ij}x_j(t) + \theta_i$.

若令 $X(t) = (x_1, \dots, x_n)^T$, $\theta = (\theta_1, \dots, \theta_n)^T$, $W =$

$(w_{ij})_{n \times n}$, 则式(1)可以写成:

$$X(t+1) = \text{sgn}(WX(t) + \theta)$$

设 $N = (W, \theta)$ 是一 DHNN, N 从初始状态 $X(t_0)$ 开始, 经过一个有限的时刻之后, 网络的输出不再发生变化, 则称此网络关于初始状态 $X(t_0)$ 稳定(或收敛). 关于任何的初态都稳定的网络称为稳定的网络. 若 X^* 是 DHNN $N = (W, \theta)$ 的一个稳定状态, 则有 $X^* = \text{sgn}(WX^* + \theta)$, 并且称 X^* 是网络 N 的稳定吸引子或简称为吸引子.

若网络每次演化都是随机等概率地选择一个神经元按式(1)进行, 并且关于任何的初态都稳定, 则称该网络是串行稳定的; 若网络每次演化所有的神经元都按式(1)进行, 并且关于任何的初态都稳定, 则称该网络是并行稳定的; 若网络每次演化都是随机等概率地选择一部分神经元“所谓的大神经元”^[6] X_1, X_2, \dots, X_m 按式(1)进行, 其中 X_1, X_2, \dots, X_m 满足 $X_j = \phi(i-j), X_1, X_2, \dots, X_m = \{1, 2, \dots, n\}$, 并且关于任何的初态都稳定, 则称该网络是部分并行稳定的. 在部分并行演化方式中, 若 X_1, X_2, \dots, X_m 满足 $X_1, X_2, \dots, X_m = \{1, 2, \dots, n\}$, 但不满足 $X_i, X_j = \phi(i-j)$, 则称这种演化方式为一部分并行演化方式, 若关于任何的初态都稳定, 则称该网络是一部分并行稳定的, 以下简称为部分并行稳定. 对于给定的网络, 它有没有稳定吸引子与网络的演化方式没有任意的关系.

3 主要结果

记 $I = \{1, 2, \dots, n\}$, $I_1 = \left\{ i \in I \mid i \geq \sum_{j=1}^n |w_{ij}|, \text{ or } i < \sum_{j=1}^n |w_{ij}|, \text{ or } w_{ii} > \sum_{j=1}^n |w_{ij}| + |i| \right\}$, $I_2 = I \setminus I_1$, $I_3 = \left\{ i \in I \mid w_{ii} < - \sum_{j=1}^n |w_{ij}| - |i| \right\}$, $W = (w_{ij})_{i,j \in I_2}$, 则有下面的几个结论.

定理 1 若存在 $\alpha_i > 0, i \in I_2$ 使矩阵 $W^* = (w_{ij}^*)_{i,j \in I_2}$ 是半正定的 (不要求对称, 下同), 则 DHNN(1) 是并行稳定的, 其中

$$w_{ij}^* = \begin{cases} iw_{ii} - \frac{1}{2} \sum_{k \in I_2} |kw_{ki} - iw_{ik}|, & \text{if } i = j \\ iw_{ij}, & \text{if } i \neq j \end{cases} \quad i, j \in I_2 \quad (2)$$

证明 由于 $x_i(t+1) = \text{sgn}(\sum_{j=1}^n w_{ij}x_j(t) + \alpha_i) = \text{sgn}(\sum_{j \in I_2} w_{ij}x_j(t) + \alpha_i)$, $i \in I_1$.

令 $\hat{\alpha}_i = \sum_{j \in I_1} w_{ij}x_j(t) + \alpha_i, i \in I_2$. 很显然, 当 $t \geq 1$ 时, I_1 中的神经元 j 的状态不会发生变化, 因此 $\hat{\alpha}_i$ 是恒定的常数, 并且 $x_i(t+1) = \text{sgn}(\sum_{j \in I_2} w_{ij}x_j(t) + \hat{\alpha}_i), i \in I_2$.

$$\text{令 } \alpha_i = \max \left\{ \sum_{j \in I_2} iw_{ij}x_j(t) + \hat{\alpha}_i \mid \sum_{j \in I_2} iw_{ij}x_j(t) + \hat{\alpha}_i < 0, x_j \in \{1, -1\}, j \in I_2 \right\} \quad (3)$$

在式(3)中, 若没有 $x_j \in \{1, -1\}, j \in I_2$, 使 $\sum_{j \in I_2} iw_{ij}x_j(t) + \hat{\alpha}_i < 0$ 成立, 则 α_i 可取为任何的负数. 设 $\bar{\alpha}_i = \hat{\alpha}_i - \frac{1}{2}, \bar{w}_{ij} = iw_{ij}, i, j \in I_2, \bar{W} = (\bar{w}_{ij})_{i,j \in I_2}$. 当 $t \geq 1$, 有

$$\begin{aligned} x_i(t+1) &= \text{sgn}(\sum_{j \in I_2} w_{ij}x_j(t) + \hat{\alpha}_i) = \text{sgn}(\sum_{j \in I_2} \bar{w}_{ij}x_j(t) + \bar{\alpha}_i) \\ &= \text{sgn}(\sum_{j \in I_2} \bar{w}_{ij}x_j(t) + \bar{\alpha}_i), i \in I_2 \end{aligned} \quad (4)$$

其中 $\bar{H}_i(t) = \sum_{j \in I_2} \bar{w}_{ij}x_j(t) + \bar{\alpha}_i, i \in I_2$, 并且当 $x_i(t) = 0, \bar{H}_i(t) = x_i(t) > 0, i \in I_2$. 很明显, DHNN(1) 的并行稳定性和 DHNN(4) 的并行稳定性相同.

定义 DHNN(4) 的能量函数为:

$$E(X(t)) = - \frac{1}{2} \sum_{i \in I_2} \sum_{j \in I_2} \bar{w}_{ij}x_i(t)x_j(t) - \sum_{i \in I_2} \bar{\alpha}_i x_i(t) \quad (5)$$

这里采用和文献[5]中, 定理 2 同样的方法 (这里的矩阵 \bar{W} 被认为是文献[5], 定理 2 中的矩阵 W) 可以证明

$$\begin{aligned} E &= E(X(t+1)) - E(X(t)) \\ &< - \frac{1}{2} \sum_{i \in I_2} \sum_{j \in I_2} w_{ij}^* x_i(t) x_j(t) \end{aligned} \quad (6)$$

其中 $I_2(t) = \{i \in I_2 \mid x_i(t) = 0\}$. 根据假设, 矩阵 W^* 是半正定的, 因此其主子矩阵也是半正定的, 所以式(6)满足

$$- \frac{1}{2} \sum_{i \in I_2(t)} \sum_{j \in I_2(t)} w_{ij}^* x_i(t) x_j(t) \leq 0$$

因此, 当 $I_2(t) \neq \emptyset$ 时, $E(t) < 0$, 即 $E(t)$ 是 DHNN(4) 的严格能量函数. 所以 DHNN(4) 是并行稳定的.

推论 1 若存在 $\alpha_i > 0, i \in I$, 使矩阵 $W^* = (w_{ij}^*)_{n \times n}$ 是半正定的, 则 DHNN(1) 是并行稳定的, 其中

$$w_{ij}^* = \begin{cases} iw_{ii} - \frac{1}{2} \sum_{k \in I} |kw_{ki} - iw_{ik}|, & \text{if } i = j \\ iw_{ij}, & \text{if } i \neq j \end{cases} \quad i, j \in I \quad (7)$$

特别地, 若 $\alpha_1 = \dots = \alpha_n = 1$, 则推论 1 与文献[5]中的定理 2 相同.

推论 2 若存在对角元素 $\alpha_i > 0, i \in I_2$ (or $i \in I$) 的对角矩阵 $\bar{W} = W$ (或 W) 是对称并且是半正定的, 则 DHNN(1) 是并行稳定的. 它是定理 1 或推论 1 的直接结果, 其中后一种情形是文献[9]中的定理 2.

例 1 设 $\alpha_i = 0, W = \begin{pmatrix} 4 & -2 & 2 \\ 1 & 3 & -2 \\ 1 & 2 & 3 \end{pmatrix}$. 则存在 $\alpha_1 = \frac{1}{2}, \alpha_2 = 1, \alpha_3 = 1$ 使对应的矩阵 $W^* = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 2 & 1 \end{pmatrix}$ 半正定, 由推论 1 知 DHNN(1) 是并行稳定的. 经计算知 16 个初始状态都并行收敛到稳定点.

定理 2 若存在 $\alpha_i > 0, i \in I_2$ (or $i \in I$) 使矩阵 $W^* = (w_{ij}^*)_{i,j \in I_2}$ (or $(w_{ij}^*)_{n \times n}$) 的主子矩阵 $W_{ii}^* (i = 1, 2, \dots, m)$ 是半正定的, 则 DHNN(1) 是部分并行稳定的, 其中 $W_{ii}^* = (w_{pq}^*)_{p,q \in X_i \in I_2}$ (or $X_i \in I$) ($i = 1, 2, \dots, m$), 而 w_{ij}^* 的表达式如(2)或式(7).

证明 按定理 1 的证明, 只需证明网络(4)是部分并行稳定的. 能量函数仍为式(5), 利用文献[5]中的定理 2 和文献[6]中的定理 1 类似的方法, 可以证明 (这里不妨假设在时刻 t 大神元 X_i 按式(1)进行演化)

$$E < - \frac{1}{2} \sum_{p \in X_i(t)} \sum_{q \in X_i(t)} w_{pq}^* x_p(t) x_q(t) \quad (8)$$

其中 $X_i(t) = \{p \in X_i \in I_2 \text{ (or } X_i) \mid x_p(t) = 0\}$. 由于 $W_{ii}^* (i = 1, 2, \dots, m)$ 半正定, 因此其主子矩阵也是半正定, 所以在式(8)中, 当 $X_i(t) \neq \emptyset$ 时, $E(t) < 0$, 即 $E(t)$ 是 DHNN(4) 的严格能量函数. 所以 DHNN(4) 是部分并行稳定的.

推论 3 若存在 $\alpha_i > 0, i \in I_2$ (or $i \in I$), 使

$$w_{ii} \geq \frac{1}{2} \sum_{j \in I_2 \text{ (or } I)} \left| w_{ij} - \frac{i}{j} w_{ji} \right|, \forall i \in I_2 \text{ (or } i \in I)$$

则 DHNN(1) 是串行稳定的. 事实上, 推论 3 是定理 2 在串行模式下的结论.

推论 4 如果存在 $\alpha_i > 0, i \in I_2$ (or $i \in I$) 使下式成立, 则 DHNN(1) 是部分并行稳定的, $w_{ii} \geq \frac{1}{2} \left(\sum_{j \in I_2 \text{ (or } I)} \left| w_{ij} - \frac{i}{j} w_{ji} \right| + \sum_{j \in X_k \in I_2 \text{ (or } X_k)} \left| w_{ij} + \frac{i}{j} w_{ji} \right| \right), i \in X_k \in I_2 \text{ (or } X_k), k = 1, 2, \dots, m$,

或 $w_{ii}^* \geq \frac{1}{2} \sum_{j \in X_k} |w_{ij} + w_{ji}|, i \in X_k, k = 1, 2, \dots, m$ (9)

显然,满足式(9)的矩阵 $W_{ii}^* = (w_{pq}^*)_{p,q \in X_i}$ ($i = 1, 2, \dots, m$) 是半正定的,据定理 2 结论成立.

例 2 设 $\theta = 0, W = \begin{pmatrix} 2 & -2 & -6 \\ 2 & 4 & 2 \\ -3 & -1 & 1 \end{pmatrix}$. 则存在 $X_1 = \{1, 2\}, X_2 = \{2, 3\}$, 则有 $W_{11}^* = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, W_{22}^* = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, 它们是半正定的矩阵. 由于推论 4 的条件(9)满足,所以网络是部分并行稳定的. 经检验知网络(1)确实是部分并行稳定的,并且稳定点是 $(-1, 1, 1)^T, (-1, -1, 1)^T, (1, 1, -1)^T$. 其它状态都部分并行收敛到这些稳定点. 当然,若 $X_1 = \{1, 2\}$ (or $\{1\}$), $X_2 = \{3\}$ (or $\{2, 3\}$), 则网络(1)同样是部分并行稳定的.

定理 3 设 $\theta = 0$. 若存在对角元素 $w_{ii} > 0$ ($i = 1, \dots, n$) 的对角矩阵 $\bar{W} = W$ 负定(不需要对称),则 DHNN(1) 没有稳定点. 特别地,若矩阵 $\bar{W} = W$ 对称且负定,则 DHNN(1) 收敛于周期为 2 的极限环.

证明 假设 DHNN(1) 有稳定点 $X = (x_1, \dots, x_n)^T$, 则 $x_i = \text{sgn}(\sum_{j=1}^n w_{ij}x_j), i = 1, \dots, n$

因此 $x_i \sum_{j=1}^n w_{ij}x_j \geq 0, x_i \sum_{j=1}^n w_{ji}x_j \geq 0, i = 1, \dots, n$. 即 $\sum_{i=1}^n \sum_{j=1}^n w_{ij}x_i x_j \geq 0$, 这与矩阵 $\bar{W} = W$ 的负定假设相矛盾. 所以 DHNN(1) 没有稳定点. 若矩阵 $\bar{W} = W$ 对称,则由文[9]的定理 3 知 DHNN(1) 将收敛到周期最多为 2 的极限环,由矩阵 \bar{W} 负定知 DHNN(1) 只能收敛到周期为 2 的极限环.

对于连接权矩阵 W 为反对称的网络,若 $w_{i1}x_1 + w_{i2}x_2 + \dots + w_{in}x_n = 0, x_j \in \{-1, 1\}, i, j = 1, 2, \dots, n$, 则反对称的 DHNN $N = (W, 0)$ 并行收敛于周期为 4 的极限环^[7,8]. 进一步,还可以得到:若存在正的对角矩阵 $\bar{W} = W$ 反对称,则满足上式的 DHNN(1) 同样也收敛于周期为 4 的极限环.

对于反对称的 DHNN,文献[8]提出猜想:若矩阵 W 反对称, $n (\geq 3), \theta = 0$ 于是 DHNN(1) 对于任何的初态 X_0 都并行收敛到周期为 4 的极限环的充分必要条件是至少存在一个神经元 $i_0 \in I$ 使得 $w_{i_0 1} + w_{i_0 2} + \dots + w_{i_0 n} = 0$ 成立.

很显然,必要条件是成立的,否则状态 $(1, 1, \dots, 1)^T$ 是稳定的. 下面说明充分条件不成立.

例 3 设 $W = \begin{pmatrix} W_{11} & 0 \\ 0 & W_{22} \end{pmatrix}, \theta = 0$, 其中 $W_{11} = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}, W_{22} = \begin{pmatrix} 0 & 1 & -2 \\ -1 & 0 & 3 \\ 2 & -3 & 0 \end{pmatrix}$. 很明显 $\sum_{j=1}^6 w_{4j} = 0$.

0,但经检验知网络(1)有周期为 4 和 12 的两种长度的极限环,因此充分条件不成立.

4 结论

文中主要研究了 DHNN 在不同的演化方式下的稳定性,并且给出了几个结果,所获结果是一些已有的稳定性结果的进一步推广,同时举例验证了结论. 很显然,文中所给的满足并行稳定条件的网络,一定是部分并行稳定(当然包括串行稳定的). 但反过来则不一定成立(例 2). 最后给出文献[8]中关于反对称 DHNN 的一个猜想的一个反例.

参考文献:

- [1] Hopfield J J. Neural network and physical systems emergent collective computational abilities [J]. Proc. Nat. Acad. Sci. USA, 1982, 79(4): 2554 - 2558.
- [2] Giles E Fogelman F, Pellegrin D. Decreasing energy functions as a tool for studying threshold Network [J]. Discrete Applied Mathematics, 1985, 12(3): 261 - 277.
- [3] Bruck J, Godman J W. A generalized convergence theorem for neural network [J]. IEEE Trans. Inform. Theory, 1988, 34(5): 1089 - 1092.
- [4] Bruck J. On the convergence properties of the Hopfield model [J]. Proc. IEEE, 1990, 78(10): 1579 - 1585.
- [5] Xu Z B, Kwong C P. Global convergence and asymptotic stability of asymmetric Hopfield neural network [J]. Journal of Mathematical Analysis and Applications, 1995, 191(3): 405 - 427.
- [6] Lee D L. New stability conditions for Hopfield neural network in partial simultaneous update mode [J]. IEEE Trans. Neural Network, 1999, 10(4): 975 - 978.
- [7] Giles E. Antisymmetrical neural network [J]. Discrete Applied Mathematics, 1986, 13(1): 97 - 100.
- [8] 许进, 保铮. 反对称离散 Hopfield 网络的稳定性理论 [J]. 电子学报, 1999, 27(1): 103 - 107.
- [9] 廖晓昕, 昌莉, 沈轶. 离散 Hopfield 神经网络的稳定性研究 [J]. 自动化学报, 1999, 25(6): 721 - 727.

作者简介:



马润年 男, 1963 年 4 月生于陕西榆林, 副教授, 博士生, 1989 年山东大学数学系运筹学与控制论专业获理学硕士学位, 现西安电子科技大学电路与系统专业博士生, 目前主要研究兴趣为最优化、神经网络、遗传算法和图论, 已发表学术论文 30 余篇.



张强 男, 1971 年生于陕西西安, 工程师, 博士生, 1999 年西安电子科技大学电路与系统专业获工学硕士学位, 现西安电子科技大学电路与系统专业博士生, 目前主要研究兴趣为神经网络、遗传算法及信号处理等, 已发表论文 10 余篇.