离散 Hopfield 神经网络的动态行为分析

李 峰, 夏靖波
(空军工程大学 电讯工程学院,陕西 西安 710077)

摘 要:离散 Hopfield 神经网络的稳定性不仅本身有重要的理论意义,而且也是网络应用的基础。主要研究非对称离散 Hopfield 神经网络在并行演化模式下的动力学行为,得到了一些新的稳定性条件,所获结果进一步推广了一些已有的结论。

关键词:神经网络;并行演化;稳定性;能量函数
中图分类号:TP183 **文献标识码**:A **文章编号**:1009-3516(2004)03-0075-03

1 基本模型

具有 n 个神经元的离散 Hopfield 神经网络记为 $N=(\boldsymbol{W},\boldsymbol{\theta})$,其演化方程为

$$x_i(t+1)=\mathrm{sgn}\left(\sum_{j=1}^{n}w_{ij}x_j(t)+\theta_i\right),\quad i=1,2,\cdots,n \tag{1}$$

其中,符号函数

$$\mathrm{sgn}(x)=\begin{cases}1 & 若\ x\geqslant 0\\ -1 & 若\ x<0\end{cases} \tag{2}$$

并且网络的连接权矩阵 $\boldsymbol{W}=(w_{ij})_{n\times n}$,阈值向量为 $\boldsymbol{\theta}=(\theta_1,\cdots,\theta_n)$,并记为 $N=(\boldsymbol{W},\boldsymbol{\theta})$,$x_i(t)$ 表示神经元 i 在时刻 t 所处的状态 $t\in\{0,1,2,\cdots\}$。时刻 t 处的状态向量为 $\boldsymbol{X}(t)=(x_1(t),\cdots,x_n(t))^{\mathrm{T}}$。在任意时刻 t ,网络所有神经元的状态都按式(1)进行演化,则称这种模式为并行(或同步)演化模式。

设网络 $N=(\boldsymbol{W},\boldsymbol{\theta})$ 从初始状态 $\boldsymbol{X}(t_0)$ 开始演化,对充分大的 t 有:

$$\boldsymbol{X}(t+1)=\boldsymbol{X}(t)$$

则称此网络关于初始状态 $\boldsymbol{X}(t_0)$ 是并行(或同步)收敛(或稳定)的。关于任何的初始状态都并行收敛的网络称为并行收敛的网络。

为叙述方便,记 $I=\{1,2,\cdots,n\}$,任给矩阵 $\boldsymbol{A}=(a_{ij})_{i,j\in I}$,记 $\boldsymbol{A}^*=(a_{ij}^*)_{i,j\in I}$,其中

$$a_{ij}^*=\begin{cases}a_{ii}-\dfrac{1}{2}\sum\limits_{k\in I}|a_{ki}-a_{ik}|, & i=j\\ a_{ij}, & i\neq j\end{cases} \tag{3}$$

2 主要结果

文献[5]中给出,如果矩阵 $\boldsymbol{A}^*=(a_{ij})_{i,j\in I}$是半正定的,并且矩阵 $\boldsymbol{B}=(b_{ij})_{i,j\in I}$是行(或列)对角占优的,即 $\forall i\in I$,有 $b_{ii}\geqslant\sum\limits_{j\neq i}|b_{ij}|$ (或 $b_{ii}\geqslant\sum\limits_{j\neq i}|b_{ji}|$),则离散 Hopfield 神经网络 $N=(\boldsymbol{A}+\boldsymbol{B},\boldsymbol{\theta})$ 是并行收敛的。这里所说的矩阵 $\boldsymbol{A}$ 半正定是指 $\forall\boldsymbol{X}\in R^n$,有 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}\geqslant 0$,矩阵 $\boldsymbol{A}$ 不需要对称. 特别地,在矩阵 $\boldsymbol{B}=\boldsymbol{0}$ 时该结论是文献[1]中定理2。进一步,在上述结论和文献[1~4]的基础上,我们给出下面的结果。

收稿日期:2004-01-05
作者简介:李 峰(1970-),男,陕西咸阳人,硕士生,主要从事信息网络技术、神经网络研究.

定理 如果矩阵 $\boldsymbol{A}^{*} = (a_{ij}^{*})_{i,j\in I}$是半正定的,矩阵 $\boldsymbol{B} = (b_{ij})_{i,j\in I}$是列对角占优的,矩阵 $\boldsymbol{C} = (c_{ij})_{i,j\in I}$是行对角占优的,则网络 $N = (\boldsymbol{A}+\boldsymbol{B}+\boldsymbol{C},\boldsymbol{\theta})$ 是并行收敛的,其中 a_{ij}^{*} 的表达见式(3)。

证明 首先令

$$\delta_i = \max\{\sum_{j=1}^{n}(a_{ij}+b_{ij}+c_{ij})x_j+\theta_i \quad \sum_{j=1}^{n}(a_{ij}+b_{ij}+c_{ij})x_j+\theta_i<0, x_j\in\{-1,1\}, j\in I\}$$

如果上式中没有 $x_1,x_2,\cdots,x_n\in\{-1,1\}$ 使 $\sum_{j=1}^{n}(a_{ij}+b_{ij}+c_{ij})x_j+\theta_i<0$ 成立,则 δ_i 可取为任何的负数。并令 $\bar{\theta}_i = \theta_i - \frac{\delta_i}{2}$,$i=1,2,\cdots,n$,于是根据符号函数的性质,有:

$$x_i(t+1) = \mathrm{sgn}(\sum_{j=1}^{n}(a_{ij}+b_{ij}+c_{ij})x_j(t)+\theta_i) = \mathrm{sgn}(\sum_{j=1}^{n}(a_{ij}+b_{ij}+c_{ij})x_j(t)+\bar{\theta}_i) \tag{4}$$

很明显,网络(4) 的收敛性与网络(1) 的收敛性完全相同,并且网络(4) 是严格的,所谓网络是严格的是指 $\forall x_j\in\{-1,1\}$,$j\in I$,有 $\sum_{j=1}^{n}(a_{ij}+b_{ij}+c_{ij})x_j+\bar{\theta}_i\neq 0$。下面将证明网络(4) 是收敛的。定义网络(4) 能量函数如下:

$$E(\boldsymbol{X}(t)) = -\frac{1}{2}\boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{A}\boldsymbol{X}(t) - \boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{B}\boldsymbol{X}(t) - \boldsymbol{X}^{\mathrm{T}}(t)\bar{\boldsymbol{\theta}} \tag{5}$$

其中 $\bar{\boldsymbol{\theta}} = (\bar{\theta}_1,\cdots,\bar{\theta}_n)^{\mathrm{T}}$。则能量函数(5) 沿网络(4) 的变化为

$$\begin{aligned}\Delta E(\boldsymbol{X}(t)) = {} & E(\boldsymbol{X}(t+1)) - E(\boldsymbol{X}(t)) = \\ & -\frac{1}{2}\boldsymbol{X}^{\mathrm{T}}(t+1)\boldsymbol{A}\boldsymbol{X}(t+1) - \boldsymbol{X}^{\mathrm{T}}(t+1)\boldsymbol{B}\boldsymbol{X}(t+1) - \boldsymbol{X}^{\mathrm{T}}(t+1)\bar{\boldsymbol{\theta}} + \\ & \frac{1}{2}\boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{A}\boldsymbol{X}(t) - \boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{B}\boldsymbol{X}(t) - \boldsymbol{X}^{\mathrm{T}}(t)\bar{\boldsymbol{\theta}} = \\ & \frac{1}{2}\Delta\boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{A}\boldsymbol{X}(t) - \frac{1}{2}\Delta\boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{A}\Delta\boldsymbol{X}(t) - \frac{1}{2}\Delta\boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{A}\Delta\boldsymbol{X}(t) - \\ & \Delta\boldsymbol{X}^{\mathrm{T}}(t)[(\boldsymbol{A}+\boldsymbol{B}+\boldsymbol{C})\boldsymbol{X}(t)+\bar{\boldsymbol{\theta}}] - \boldsymbol{X}^{\mathrm{T}}(t+1)\boldsymbol{B}\Delta\boldsymbol{X}(t) + \Delta\boldsymbol{X}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{X}(t) = \\ & -p(t) - q(t) - r(t)\end{aligned} \tag{6}$$

其中

$$\begin{aligned}\Delta\boldsymbol{X}(t) &= \boldsymbol{X}(t+1) - \boldsymbol{X}(t) \\ p(t) &= -\frac{1}{2}\Delta\boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{A}\boldsymbol{X}(t) + \frac{1}{2}\boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{A}\Delta\boldsymbol{X}(t) + \frac{1}{2}\Delta\boldsymbol{X}^{\mathrm{T}}(t)\boldsymbol{A}\Delta\boldsymbol{X}(t) \\ q(t) &= \Delta\boldsymbol{X}^{\mathrm{T}}(t)[(\boldsymbol{A}+\boldsymbol{B}+\boldsymbol{C})\boldsymbol{X}(t)+\bar{\boldsymbol{\theta}}], \\ r(t) &= \boldsymbol{X}^{\mathrm{T}}(t+1)\boldsymbol{B}\Delta\boldsymbol{X}(t) - \Delta\boldsymbol{X}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{X}(t)\end{aligned}$$

很显然,当 $\Delta x_i(t)\neq 0$ 时,$\Delta x_i(t) = 2x_i(t+1) = -2x_i$。由于网络(4) 是严格的,因此,容易证明 $q(t)\geqslant 0$,并且当 $\Delta X\neq 0$ 时,$q(t)\geqslant 0$。

在并行演化模式下,记

$$I_2(t) = \{i\in I \mid \Delta x_i(t)\neq 0\}, I_1(t) = I/I_2(t)$$

利用矩阵 $\boldsymbol{B}$ 是列对角占优的、矩阵 $\boldsymbol{C}$ 是行对角占优的特点,可以得到

$$\begin{aligned}r(t) = {} & \boldsymbol{X}^{\mathrm{T}}(t+1)\boldsymbol{B}[\boldsymbol{X}(t+1)-\boldsymbol{X}(t)] - [\boldsymbol{X}^{\mathrm{T}}(t+1)-\boldsymbol{X}^{\mathrm{T}}(t)]\boldsymbol{C}\boldsymbol{X}(t) = \\ & [\boldsymbol{X}^{\mathrm{T}}(t+1)-\boldsymbol{X}^{\mathrm{T}}(t)]\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}(t+1) - [\boldsymbol{X}^{\mathrm{T}}(t+1)-\boldsymbol{X}^{\mathrm{T}}(t)]\boldsymbol{C}\boldsymbol{X}(t) = \\ & \sum_{i\in I}[x_i(t+1)-x_i(t)]\sum_{j\in I}b_{ji}x_j(t+1) - \sum_{i\in I}[x_i(t+1)-x_i(t)]\sum_{j\in I}c_{ij}x_j(t) = \\ & 2\sum_{i\in I_2(t)}x_i(t+1)\sum_{j\in I}b_{ji}x_j(t+1) + 2\sum_{i\in I_2(t)}x_i(t)\sum_{j\in I}c_{ij}x_j(t) \geqslant \\ & 2\sum_{i\in I_2(t)}(b_{ii}-\sum_{j\neq i}|b_{ji}|) + 2\sum_{i\in I_2(t)}(c_{ii}-\sum_{j\neq i}|c_{ji}|) \geqslant 0\end{aligned}$$

利用文献[1] 中定理 2 完全同样的方法可以证明

$$p(t) = \frac{1}{2}\sum_{i\in I_2(t)}\Delta x_i(t)[\sum_{j\in I_2(t)}a_{ij}\Delta x_i(t) - \sum_{j\in I}(a_{ij}-a_{ji})x_j(t)] = \frac{1}{2}\sum_{i\in I_2(t)}\sum_{j\in I_2(t)}\Delta x_i(t)a_{ij}^{*}\Delta x_i(t)$$

很显然由于矩阵 $A^* = (a_{ij}^*)_{i,j\in I}$ 是半正定的,所以其主子式 $(a_{ij}^*)_{i,j\in I_2(t)}$ 也是半正定的。因此, $p(t) \geqslant 0$。

综合上述知,式(6) 可以化为 $\Delta E(X(t)) \leqslant 0$,并且当 $\Delta X(t) \neq 0$ 时, $\Delta E(X(t)) < 0$。所以能量函数 $E(X(t))$ 是严格的能量函数,所以定理成立。

说明:在定理中,当矩阵 $B = C = 0$ 时,定理就是文献[1] 中的定理 2;当矩阵 $B = 0$ 或 $C = 0$ 时就是文献[4 - 5] 中定理的特殊情况. 并且对于特殊情况的矩阵 A,我们还可以给出下面的推论。

推论　如果矩阵 A 是对称的并且是半正定的,矩阵 $B = (b_{ij})_{i,j\in I}$ 是列对角占优的,矩阵 $C = (c_{ij})_{i,j\in I}$ 是行对角占优的,则网络 $N = (A + B + C, \theta)$ 是并行收敛的. 特别地,如果矩阵 $A = 0$,并且矩阵 $B = (b_{ij})_{i,j\in I}$ 是列对角占优的,矩阵 $C = (c_{ij})_{i,j\in I}$ 是行对角占优的,则网络 $N = (B + C, \theta)$ 是并行收敛的。

例子　考虑网络 $N = (W, 0)$ 的并行收敛性,其中矩阵 $W = \begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix}$。显然, $W^* = \begin{pmatrix} 1.5 & 4 \\ 2 & 4.5 \end{pmatrix}$ 并不是半正定的,无法使用文献[2] 中定理 2 的结论。这样一个非常简单的 2 个神经元的离散 Hopfield 神经网络,也很难用已有的结果直接给出结论。但是,利用本文的推论,很易得出网络是并行收敛的。这是因为 $W = B + C$,其中 $B = \begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ 分别是列对角占优的和行对角占优的。事实上,经计算知,网络有 2 个稳定状态 $X_0 = (-1, -1)^T$, $X_2 = (1, 1)^T$ 并且状态 $X_1 = (1, -1)^T$, $X_2 = (-1, 1)^T$ 分别收敛到 X_0 和 X_3。

参考文献:

[1] Xu Z B, Kwong C P. Global Convergence and Asymptotic Stability of Asymmetric Hopfield Neural Networks[J]. Mathematical Analysis and Applications, 1995, 191(3): 405 - 427.

[2] Lee D L. New Stability Condotions for Hopfield Neural Networks in Partial Simultaneous Update Mode[J]. IEEE Trans. Neural Networks, 1999, 10(4): 975 - 978.

[3] 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.

[4] 马润年, 杨友社. 离散 Hopfield 网络的稳定性[J]. 空军工程大学学报(自然科学版), 2001, 1(4): 39 - 41.

[5] 马润年, 张　强, 许　进.. 不同演化模式的离散 Hopfield 网络的稳定性研究[J]. 系统工程与电子技术, 2002, 24(3): 90 - 94.

(编辑:门向生)

Dynamic Behavior Analysis of Discrete Hopfield Neural Networks

LI Feng,　XIA Jing - bo

(The Telecommunication Engineering Institute, Air Force Engineering University, Xi'an, Shaanxi 710077, China)

Abstract: The stability of discrete Hopfield neural networks not only has an important theoretical significance, but also is the foundation of the applications of the networks. In this paper, the dynamic behavior of asymmetric discrete Hopfield neural network is studied mainly in parallel updating mode, and some new stability conditions of the neural networks are given. The obtained results serve the function of further generalizing some existing conclusions.

Key words: neural networks; parallel updating; stability; energy function