第 30 卷 第 4 期
2004 年 7 月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 30, No. 4
July, 2004

Hopfield-型网络求解优化问题的一般演化规则[1)]

邱深山 邓飞其 刘永清
(华南理工大学自动控制与工程系 广州 510640)
(E-mail: ssqiu@21cn.com)

摘 要 基于离散 Hopfield-型网络和延迟离散 Hopfield-型网络求解优化问题提出了两种一般演化规则,演化序列的动态阈值是这些规则的重要特征,并获得了收敛性定理. 推广了已有的离散 Hopfield-型网络和延迟离散 Hopfield-型网络的收敛性结果,给出了能量函数局部极大值点与延迟离散 Hopfield-型网络的稳定态的关系的充分必要条件. 鉴于延迟离散 Hopfield-型网络更有效地应用于优化计算问题,给出了一般分解策略. 实验表明与离散 Hopfield-型网络的算法相比,文中提出的算法既有较高的收敛率又缩短了演化时间.

关键词 离散 Hopfield-型网络,延迟,收敛性,稳定态
中图分类号 TP18; TP311

A Generalized Updating Rules Using Hopfield-Type Neural Networks for Optimization Problems

QIU Shen-Shan DENG Fei-Qi LIU Yong-Qing
(*Department of Automatic & Engineering, South China University of Technology, Guangzhou* 510640)
(E-mail: ssqiu@21cn.com)

Abstract This paper presents two generalized updating rules based on Hopfield-type neural networks (with delay or without delay) for optimization problems. These rules are characterized by dynamic thresholds of the updating sequence. Convergence theorems of discrete Hopfield-type neural networks with delay are obtained, which extend the exsiting convergence results. Also obtained is a sufficient and necessary condition for the relation between the stable states of neural networks and the points of local maximum value of energy function. Decomposed strategy is given in order to apply the Hopfield-type neural networks with delay to optimization problems effectively. Finally, the experimental results demonstrate that the given algorithm improves the convergence rate and decreases the updating time when compared with Hopfield-type neural network without delay.

Key words Discrete Hopfield-type neural network, delay, convergence, stable state

1) 国家自然科学基金(69934030, 69874015, 60374023)和华南理工大学自然科学基金资助
Supported by National Natural Science Foundation of P. R. China(69934030, 69874015, 60374023) and the Natural Science Foundation of South China University of Technology
收稿日期 2002-07-03 收修改稿日期 2002-11-27
Received July 3, 2002; in revised form November 27, 2002

1 引言

Hopfield-型神经网络最典型的应用是NP-完全问题的一种有效算法[1,2]. 特别是离散型神经网络不但远比连续型网络更有效,而且更易于工程实现[3]. 然而,离散Hopfield-型网络用于求解优化问题理论基础仍然存在局限性:陷入局部极值问题,权阵对角元非负是收敛性的必要条件之一[1,3,4]. 虽然,Michel[4]和Xu[5]将其进一步推广,但权阵对角元非负的限制仍然是收敛性的必要条件之一. 能否突破这一限制,是本文研究的主要内容.

本文没有拘泥于原来的神经元状态输出函数sgn(·)形式的限制,提出了两种一般演化规则,收敛的必要条件可以不必限制"网络连接权阵的对角元非负",不仅给出了新的收敛性定理,还给出了能量函数的极大值点与网络稳定态的关系,使联想记忆设计上更加灵活、应用范围更加广泛,拓宽了离散Hopfield-型网络用于求解优化问题的适用范围.

2 离散Hopfield-型网络及其相关的定义

延迟离散Hopfield-型网络[5~8]可以表示为

$$x_i(t+\tau)=\sigma_i(H_i(\boldsymbol{x}(t)))\in\{-1,1\} \tag{1}$$

其中 $H_i(\boldsymbol{x}(t))=\sum_{j=1}^{n}\sum_{l=0}^{k}w_{ij}^{l}x_j(t-l\tau)-\theta_i$,$\sigma_i(\cdot)$是第$i$个神经元的输出函数.

一般来说,τ和k由具体应用问题而定. 延迟离散Hopfield-型神经网络中具有典型意义和演化特征的是$k=1$时的延迟离散Hopfield-型网络,简记$N=(w^0\oplus w^1,\boldsymbol{\theta})$,有时称之为二阶延迟离散Hopfield-型网络. 当取$\sigma_i(\cdot)=\mathrm{sgn}(\cdot)$时,我们曾给出了一些结果[7~9]. 当二阶延迟离散Hopfield-型网络$N=(w^0\oplus w^1,\boldsymbol{\theta})$的延迟权阵$w^1=O_{n\times n}$并取$\sigma_i(\cdot)=\mathrm{sgn}(\cdot)$时,它为离散Hopfield-型网络[1,3~5].

用B^n表示每个分量仅取±1的n维向量全体,即$B^n=\{\boldsymbol{v}:\boldsymbol{v}=(v_1,v_2,\cdots v_n)^{\mathrm{T}},v_i\in\{-1,1\},i=1,2,\cdots,n\}$,$\langle\boldsymbol{u},\boldsymbol{v}\rangle$表示向量$\boldsymbol{u},\boldsymbol{v}\in B^n$的内积,即$\langle\boldsymbol{u},\boldsymbol{v}\rangle=\sum_{i=1}^{n}u_iv_i$. $d_H(\boldsymbol{u},\boldsymbol{v})$表示$\boldsymbol{u}$与$\boldsymbol{v}$的Hamming距离. 显然有$\langle\boldsymbol{u},\boldsymbol{v}\rangle=n-2d_H(\boldsymbol{u},\boldsymbol{v})$. 用$B_H(\boldsymbol{v},r)$表示$B^n$中与$\boldsymbol{v}$的Hamming距离不超过$r$的向量全体,即$B_H(\boldsymbol{v},r)=\{\boldsymbol{u}:d_H(\boldsymbol{u},\boldsymbol{v})\leqslant r;\boldsymbol{u},\boldsymbol{v}\in B^n$.

定义1. 二阶延迟离散Hopfield-型网络$N=(w^0\oplus w^1,\theta)$的一个状态$\boldsymbol{v}^*\in B^n$称为稳定态(或称不动点),任意$i,1\leqslant i\leqslant n$有

$$v_i^*=f\left(\sum_{j=1}^{n}w_{ij}^0v_j^*+\sum_{j=1}^{n}w_{ij}^1v_j^*-\theta_i\right) \tag{2}$$

成立,其中$\boldsymbol{v}^*=(v_1^*,v_2^*,\cdots,v_n^*)^{\mathrm{T}}$,$f$是与演化规则相关联的输出函数.

定义2[7~9]. B^n上二元向量函数$E(\boldsymbol{u},\boldsymbol{v})=\boldsymbol{u}^{\mathrm{T}}w^0\boldsymbol{u}+2\boldsymbol{u}^{\mathrm{T}}w^1\boldsymbol{v}-2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\theta}$称为网络$N=(w^0\oplus w^1,\boldsymbol{\theta})$的能量函数,其中$\boldsymbol{u}=\boldsymbol{v}(t)$,$\boldsymbol{u}=\boldsymbol{v}(t-1)$,简记$E(t)\equiv E(\boldsymbol{v}(t),\boldsymbol{v}(t-1))$.

3 优化问题和一般演化规则

用离散Hopfield-型网络求解优化计算问题的第一步是将其转化为如下形式的最大值

或最小值问题

$$E(\boldsymbol{v},\alpha,\beta)=\alpha\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}\,v_iv_j-\beta\sum_{i=1}^{n}v_i\theta_i \tag{3}$$

Shrivastava 和 Michel 早期研究的是 $\alpha=-\frac{1}{2},\beta=-1$ 时的能量函数形式[1,4]；而 Vidyanagar[10] 研究了 $\alpha=\frac{1}{2},\beta=-1$ 时的能量函数形式.

当然,还有许多优化问题可以通过适当的变换转化为式(3)问题,如下二次优化问题

$$\min_{x\in B_0^n}(\|A\boldsymbol{v}-\boldsymbol{y}\|^2+\eta\|D\boldsymbol{v}\|^2) \tag{4}$$

可以转化为式(3)(当 $\alpha=-\frac{1}{2},\beta=-1$)形式的最小值问题. 其中 $w=(w_{ij})_{n\times n}=-A^{\mathrm T}A-\lambda D^{\mathrm T}D$, $\boldsymbol{\theta}=-A^{\mathrm T}\boldsymbol{y}$, $B_0^n\equiv\{\boldsymbol{v}=(v_1,v_2,\cdots,v_n)^{\mathrm T}\mid v_i=\pm1(0,1),i=1,2,\cdots,n\}$,$\|\cdot\|$ 是 l_2 范数. 如计算机视觉、模式识别和图论等领域的组合优化问题均可转化为式(3)求解,Ramanujam[3] 给出了部分应用问题的映射关系. 众所周知,用离散 Hopfield-型网络求解优化计算问题(3)的理论基础是离散 Hopfield-型网络拥有收敛性,但收敛性的条件要求权阵必需满足对称和对角元非负[1,3~5],或伪对称对角元非负[4]；若是非对称,仍然需对角元非负[5]. 对于式(3)一般的权阵 w,能否用离散 Hopfield-型网络求解优化计算问题(3)还是一个公开问题. 为了使离散 Hopfield-型网络能求解更广泛的优化计算问题,将其演化规则推广到一般情况,现详述如下.

定义 3. $N=(w^0\oplus w^1,\boldsymbol{\theta})$ 表示二阶延迟离散 Hopfield-型网络,任选 $\boldsymbol{v}(0)=\boldsymbol{v}(1)\in B^n$ 为初值,任意选取神经元的演化序列 $L(k)\subseteq\{1,2,\cdots,n\}$($L(k)$ 表示神经元的演化序列指标集). 对 $k\geqslant1$,如果 $i\notin L(k)$,令 $v_i(k+1)=v_i(k)$；否则 $i\in L(k)$,依如下演化规则进行：

$$v_i(k+1)=f(\boldsymbol{v}(k),\boldsymbol{v}(k-1))\equiv\begin{cases}-v_i(k), & v_i(k)H_i(k)<-T_i(k)\\ v_i(k), & \text{否则}\end{cases} \tag{5}$$

我们称其为一般演化规则,或者简称为 GURD1 规则(General Updating Rule with Delay),其中 $H_i(k)=\sum_{j=1}^{n}(w_{ij}^0v_j(k)+w_{ij}^1v_j(k-1)-\theta_i)$, $T_i(k)\equiv\sum_{j\in L(k)}|w_{ij}^0|+2\sum_{j\in L(k-1)}|w_{ij}^1|$.

由以上定义,文献[6]的一般演化规则是延迟项 $w^1=O_{n\times n}$ 时的特例. 进一步将定义修改为如下规则,其优点是展示了 w^0 的主对角元的作用. 今后讨论的演化规则均以定义 4 为基础,除非特殊说明.

定义 4. 如果在定义 3 中取 $T_i(k)\equiv-w_{ii}^0+\sum_{j\in L(k),j\neq i}|w_{ij}^0|+2\sum_{j\in L(k-1)}|w_{ij}^1|$ 时,我们称其为一般演化规则,并简称为 GURD2 规则.

如上定义中值得注意的是均假定 w^0 是对称的. 其实,将延迟网络应用到类似于式(3)的优化问题中,讨论 w^0 非对称性情况似乎意义不大,因为有"任何非对称矩阵均可以分解成对称矩阵和反对称矩阵之和",使式(3)的非对称优化问题分解成与延迟网络相对应的问题成为可能.

一般演化规则的提出是使神经网络可以求解更广泛的优化问题. 也就是说,在一般演化规则下,网络的稳定点恰好是优化问题的一个可行解(一般来说是优化问题局部极值点). 当然,定义 2 给出的能量函数是神经网络和优化问题的联系桥梁. 如下的例子说明了一般

演化规则的演化过程,演化具体过程见表 1 和表 2.

表 1 基于异步一般演化规则,对所有的初值的演化过程

编码	状态向量编码/神经元演化指标								稳定态	与稳定态对应的能量函数值
	1	2	3	4	1	2	3	4		
0→	8	12	12	13	13	13	13	13	13	次最大值
1→	9	13	13	13	13	13	13	13	13	次最大值
2→	10	14	14	15	15	15	15	15	15	最大值
3→	11	15	15	15	15	15	15	15	15	最大值
4→	4	4	4	5	5	5	5	5	5	→C
5→	5	5	5	5	5	5	5	5	5	C
6→	6	6	6	6	6	6	6	6	6	B
7→	7	7	7	7	7	7	7	7	7	A
8→	8	12	12	13	13	13	13	13	13	次最大值
9→	9	13	13	13	13	13	13	13	13	次最大值
10→	10	14	14	15	15	15	15	15	15	最大值
11→	11	15	15	15	15	15	15	15	15	最大值
12→	12	12	12	13	13	13	13	13	13	次最大值
13→	13	13	13	13	13	13	13	13	13*	次最大值
14→	14	14	14	15	15	15	15	15	15	最大值
15→	15	15	15	15	15	15	15	15	15*	最大值

表 2 基于异步一般演化规则对应所有初值的能量函数值和稳定态

编码	对应初值的能量函数值(EFValue)	稳定态 (SP)	与稳定态对应的能量函数值
0	−49.8824	13	次最大值
1	−56.5568	13	次最大值
2	31.4276	15	最大值
3	−3.24682	15	最大值
4	71.9292	5	→C
5	113.255	5	C
6	121.239	6	B
7	134.565	7	A
8	5.43522	13	次最大值
9	30.7608	13	次最大值
10	70.7452	15	最大值
11	68.0708	15	最大值
12	103.247	13	次最大值
13	176.572	13	次最大值
14	136.557	15	最大值
15	181.882	15	最大值

例 1. 二阶延迟离散 Hopfield-型网络 $N=(w^0\oplus w^1,\boldsymbol{\theta})$ 的具体形式如下,其神经元演化序列为 $L(k)\subseteq\{1,2,3,4\}L(k)=\{l\},l=k-[(k-1)/4]\times 4,k\geqslant 1$. $[s]$表示不超过 s 的最大整数.

$$w^0=\begin{bmatrix}-7 & -2 & -2 & 4\\ -2 & -5 & -4 & 6\\ -2 & -4 & 11 & -3.5\\ 4 & 6 & -3.5 & 0\end{bmatrix},\quad w^1=\begin{bmatrix}10 & -2 & 2 & -4\\ 1 & 6 & 2.5 & -1\\ -2 & -2.5 & 10 & 5\\ 4 & 1 & -5 & 10\end{bmatrix},\quad \theta=\begin{bmatrix}-12.8294\\ -29.4529\\ -10.8275\\ -4.83139\end{bmatrix}$$

例 1 可以具体再现一般演化规则的动态过程,它根据所选取神经元的演化序列 $L(k)\subseteq\{1,2,\cdots,n\}$ 的不同而变化. 选择神经元的演化序列尽管有无穷多种方式,人们已经注意到了与优化问题对应的能量函数的最大(小)值点,不会因演化序列的不同而变化的. 鉴于演

化序列的复杂性，许多研究者还是集中选择容易处理的演化模式：同步、异步演化和部分同步的演化模式进行研究[1~13]. 无论选择什么样的演化模式，其共同的目标是保持能量函数的单调性(单调递增或递减)，进而获得收敛性定理. 以能量函数单调变化速度为评价标准，选择什么样的演化序列是最好(能逃离局部最优而达到全局最优的策略)仍然是一个公开问题.

虽然 Hopfield-型网络在异步演化模式下是收敛的，但存在两个问题，一是对网络连接权阵的限制，二是 Hopfield-型网络稳定滞后于相应的能量函数的收敛的步数最多达 n^2[11]. 定义 3 和定义 4 的规则与 Hopfield-型网络的异步演化规则不同，前者的演化规则是根据所选取神经元的演化序列的不同变化的，特别其输出函数不是固定的.

4 基于 GURD 规则的收敛性定理

基于一般演化规则，以下给出二阶延迟离散 Hopfield-型网络 $N=(w^0\oplus w^1,\boldsymbol{\theta})$ 的收敛性定理.

4.1 主要定理的证明和引理

引理 1[8]. 二阶延迟离散 Hopfield-型网络 $N=(w^0\oplus w^1,\boldsymbol{\theta})$，$w^0$ 是 $n\times n$ 阶对称矩阵，w^1 是 $n\times n$ 阶矩阵且对角元素满足

$$w_{ii}^0 \geqslant 0,\quad w_{ii}^1 \geqslant \sum_{\substack{j=1\\ j\neq i}}^{n} |w_{ji}^1|,\ i=1,2,\cdots,n \tag{6}$$

则网络从任意的初始状态 $\boldsymbol{x}(0)=\boldsymbol{x}(1)\in B^n$ 异步方式运行，总能收敛到一个稳定态.

定理 1. 二阶延迟离散 Hopfield-型网络 $N=(w^0\oplus w^1,\boldsymbol{\theta})$，$w^0$ 是 $n\times n$ 阶对称矩阵，w^1 是 $n\times n$ 阶矩阵且对角元素满足 $w_{ii}^1\geqslant\sum\limits_{\substack{j=1\\ j\neq i}}^{n}|w_{ji}^1|$，则对于任意选取的神经元的演化序列 $L(k)\subseteq\{1,2,\cdots,n\}$，依 GURD1 演化规则进行演化，$k\geqslant1$，当 $\Delta\boldsymbol{x}(k)\neq0$ 或 $\Delta\boldsymbol{x}(k-1)\neq0$，有

$$\Delta E(k)\equiv E(\boldsymbol{x}(k+1),\boldsymbol{x}(k))-E(\boldsymbol{x}(k),\boldsymbol{x}(k-1))\geqslant 0 \tag{7}$$

成立；当 $\Delta\boldsymbol{x}(k)\neq0$ 有 $\Delta E(k)>0$，网络最终达到稳定态，即存在 $K_0>0$，当 $k\geqslant K_0$ 有 $\boldsymbol{x}(k+1)=\boldsymbol{x}(k)=\boldsymbol{x}(k-1)$. 也就是说网络从任意的初始状态 $\boldsymbol{x}(0)=\boldsymbol{x}(1)\in B^n$ 依 GURD1 演化规则进行演化，总能收敛到网络的一个稳定态.

证明. 由定义 2 给出的能量函数

$$E(\boldsymbol{u},\boldsymbol{v}) = \boldsymbol{u}^{\mathrm{T}}w^0\boldsymbol{u}+2\boldsymbol{u}^{\mathrm{T}}w^1\boldsymbol{v}-2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\theta}$$

其中 $\boldsymbol{u}=\boldsymbol{x}(k)$，$\boldsymbol{v}=\boldsymbol{x}(k-1)$，$\Delta\mathrm{E}(k)\equiv\mathrm{E}(k+1)-\mathrm{E}(k)$，$\Delta x_i(k)\equiv x_i(k+1)-x_i(k)$，$\theta$ 为阈值向量 $\boldsymbol{\theta}=(\theta_1,\theta_2,\cdots,\theta_n)^{\mathrm{T}}$. 显然 $\mathrm{E}(k)\equiv\mathrm{E}(\boldsymbol{x}(k),\boldsymbol{x}(k-1))$ 有上界，即

$$|E(k)|\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}(|w_{ij}^0|+2|w_{ij}^1|)+2\sum_{i=1}^{n}|\theta_i|\equiv M \tag{8}$$

则

$$\begin{aligned}\Delta\mathrm{E}(k)\equiv{}&\boldsymbol{x}^{\mathrm{T}}(k+1)w^0\boldsymbol{x}(k+1)+2\boldsymbol{x}^{\mathrm{T}}(k+1)w^1\boldsymbol{x}(k)-2\boldsymbol{x}^{\mathrm{T}}(k+1)\boldsymbol{\theta}-\\&\boldsymbol{x}^{\mathrm{T}}(k)w^0\boldsymbol{x}(k)-2\boldsymbol{x}^{\mathrm{T}}(k)w^1\boldsymbol{x}(k-1)+2\boldsymbol{x}^{\mathrm{T}}(k)\boldsymbol{\theta}\end{aligned}$$

对于 $k\geqslant2$，$i\in L(k)\subseteq\{1,2,\cdots,n\}$，依 GURD1 演化规则进行演化，则有

$$\begin{aligned}\Delta E(k)=&\sum_{i=1}^{n}2\Delta x_i(k)H_i(\boldsymbol{x}(k))+\sum_{i=1}^{n}\sum_{j=1}^{n}(\Delta x_i(k))(\Delta x_j(k))w_{ij}^0+\\&2\sum_{i=1}^{n}\sum_{j=1}^{n}x_i(k+1)\Delta x_j(k-1)w_{ij}^1=\\&\sum_{i=1}^{n}2\Delta x_i(k)\left\{H_i(\boldsymbol{x}(k))+\frac{1}{2}\sum_{j=1}^{n}(\Delta x_j(k))w_{ij}^0+\sum_{j=1}^{n}\Delta x_j(k-1)w_{ij}^1\right\}+\\&2\sum_{j\in L(k-1)}\Delta x_j(k-1)\left\{\sum_{i=1}^{n}x_i(k)w_{ij}^1\right\}=\\&\sum_{i\in L(k)}2\Delta x_i(k)\left\{H_i(\boldsymbol{x}(k))+\frac{1}{2}\sum_{j\in L(k)}(\Delta x_j(k))w_{ij}^0+\sum_{j\in L(k-1)}(\Delta x_j(k-1))w_{ij}^1\right\}+\\&2\sum_{i\in L(k-1)}\Delta x_i(k-1)\left\{\sum_{j=1}^{n}x_j(k)w_{ji}^1\right\}\end{aligned}\tag{9}$$

注意到式(9)最后的等式中的第二项,由 $w_{ii}^1\geqslant\sum_{\substack{j=1\\j\neq i}}^{n}|w_{ji}^1|$,有 $\sum_{i\in L(k-1)}\Delta x_i(k-1)\left\{\sum_{j=1}^{n}x_j(k)w_{ji}^1\right\}\geqslant 0$. 以下考虑第一项:

i)对 $i\in L(k)$,依 GURD1 演化规则,当 $x_i(k)=-1,H_i(k)>\sum_{j\in L(k)}|w_{ij}^0|+2\sum_{j\in L(k-1)}|w_{ij}^1|$ 时,有 $x_i(k+1)=1$,则 $(\Delta E(k))_i>0$;

ii)对 $i\in L(k)$,依 GURD1 演化规则,当 $x_i(k)=1,H_i(k)<-\sum_{j\in L(k)}|w_{ij}^0|-2\sum_{j\in L(k-1)}|w_{ij}^1|$ 时,有 $x_i(k+1)=-1$,则 $(\Delta E(k))_i>0$;

iii)对 $i\in L(k)$,依 GURD1 演化规则,当 $\Delta x_i(k)=0,\Delta x_i(k-1)\neq 0$ 时,由于 $w_{ii}^1\geqslant\sum_{\substack{j=1\\j\neq i}}^{n}|w_{ji}^1|$,则 $(\Delta E(k))_i\geqslant 0$. 综合以上结果可以得到,对于 $\Delta\boldsymbol{x}(k)\neq 0$ 或 $\Delta\boldsymbol{x}(k-1)\neq 0$ 时,有 $\Delta E(k)\geqslant 0$.

以下证明其最终收敛到网络的稳定态. 假设网络不收敛于稳定态,也就是说,存在一个序列 $\{k_0,k_1,k_2,\cdots,\}$, $k_0\geqslant 1$,满足 $k_{i+1}-k_i>3$(由于 $\{k_i\}$ 是无穷序列),使得 $\boldsymbol{x}(k_m+1)\neq\boldsymbol{x}(k_m)$(或 $\boldsymbol{x}(k_m)\neq\boldsymbol{x}(k_m-1)$ 这种情况可以不必考虑)和 $\boldsymbol{x}(k_m+1)=\boldsymbol{x}(k_m+2)=\boldsymbol{x}(k_m+3)=\cdots=\boldsymbol{x}(k_{m+1})$, $m=0,1,\cdots$. 如果 $\boldsymbol{x}(k_m+1)\neq\boldsymbol{x}(k_m)$,由上面的证明 i), ii) 有 $\Delta E(k_m)>0$;如果 $\boldsymbol{x}(k_m+i)=\boldsymbol{x}(k_m+i+1)=\boldsymbol{x}(k_m+i+2)$,有 $\Delta E(k_m+i)=0$. 所以

$$\begin{aligned}E(k_m+1)\equiv E(k_m)+\Delta E(k_m)=&E(k_{m-1}+1)+\Delta E(k_m)=\\&E(k_{m-1})+\Delta E(k_{m-1})+\Delta E(k_m)=E(k_0)+\sum_{j=0}^{m}\Delta E(k_j)\end{aligned}$$

显然,结合式(8)并对上式取 $m\to+\infty$ 的极限,有

$$\begin{aligned}2\left(\sum_{i=1}^{n}\sum_{j=1}^{n}(|w_{ij}^0|+2|w_{ij}^1|)+2\sum_{i=1}^{n}|\theta_i|\right)\geqslant&\lim_{m\to+\infty}\{E(k_m+1)-E(k_0)\}=\\&\lim_{m\to+\infty}\sum_{j=0}^{m}\Delta E(k_j)=+\infty\end{aligned}\tag{10}$$

现在说明式(10)的右端的极限值为什么是 $+\infty$,由于 $\boldsymbol{x}$ 的取值最多为 2^n 个不同的值,故 $\Delta E(\boldsymbol{x}(k))$ 最多为 $2^{2n}(2^{2n}-1)+1$ 个不同的值,所以式(10)的右端极限值是 $+\infty$ 与式(10)的左端产生矛盾. 所以存在 $K_0>0$,当 $k\geqslant K_0$ 有 $\boldsymbol{x}(k+1)=\boldsymbol{x}(k)=\boldsymbol{x}(k-1)$,即网络最终收敛到

稳定态.　　证毕.

定理 2. 在定理 1 的假定条件下，网络从任意的初始状态 $\boldsymbol{x}(0)=\boldsymbol{x}(1)\in B^n$ 依 GURD2 演化规则进行演化，总能收敛到一个网络的稳定态.

定理 2 的证明与定理 1 类似，略去.

由表 1 和 2 中可以看出一般演化规则的收敛速度[16,17]和收敛到能量函数最大值的可能性分布，即便不是最大值也是次最大值，除此之外的稳定态也是能量函数的相对最大值点. 而离散 Hopfield 网络算法异步演化模式，随机选择演化神经元演化常常不能在每一步保证使能量函数递增.

引理 1 中对 w^0 主对角元 $\{w_{ii}^0: i=1,2,\cdots,n\}$ 要求非负[15,16]，但在一般演化规则下，定理 1 和定理 2 表明，可以去掉这一限制. 对 w^1 仍然要求列对角占优的.

4.2　网络的稳定态与 GURD 演化规则的关系

定理 1 和 2 说明网络能收敛到稳定态，但没有给出有关稳定态可用的信息，也就是说，当网络稳定时，稳定态应该满足什么条件？以下分析稳定态的第 i 个分量 $x_i(k)$ 与 $H_i(k)$ 的限定条件 $H_i(k)>T_i(k)$ 或 $H_i(k)<-T_i(k)$ 的关系.

定理 3. 二阶延迟离散 Hopfield-型网络 $N=(w^0\oplus w^1,\boldsymbol{\theta})$，$w^0$ 是 $n\times n$ 阶对称矩阵，w^1 是 $n\times n$ 阶矩阵且对角元素满足 $w_{ii}^1\geqslant\sum_{\substack{j=1\\j\neq i}}^{n}|w_{ji}^1|$，则对于任意选取的神经元的演化序列 $L(k)\subseteq\{1,2,\cdots,n\}$，依 GURD1(2) 演化规则进行演化，$\boldsymbol{x}(k)=\boldsymbol{x}^*\ (k>k_0)$ 是稳定态的充分必要条件：

1) 如果 $\boldsymbol{x}(k)$ 的第 i 个分量的对应阈值满足 $x_i(k)=1$ 时，$H_i(k)>T_i(k)$；

2) 如果 $\boldsymbol{x}(k)$ 的第 i 个分量的对应阈值满足 $x_i(k)=-1$ 时，$H_i(k)<-T_i(k)$；

3) $\bigcup_{k>k_0}L(k)=\{1,2,\cdots,n\}$.

证明. 充分条件. 显然.

必要条件：如果 $k>k_0$，$x(k)=x^*$ 是稳定态，由定义 3 的式(5)的演化规则，不会出现 $x_i(k+1)=-x_i(k)$ 的情况，所以条件 1) 和 2) 满足，同时对所有的神经元成立. 也就是说，当 $k>k_0$ 时，演化序列中最终能将所有的神经元遍历到，所以条件 3) 成立.　　证毕.

4.3　离散 Hopfield-型网络与延迟离散 Hopfield-型网络的关系

如果分解 w 为 $w=w^0+2w^1$，其中 $w^0=\frac{1}{2}(w^{\mathrm{T}}+w)-\mathrm{diag}(w)+\mathrm{diag}(\alpha)$，$w^1=\left(\frac{1}{4}(w-w^{\mathrm{T}})+\frac{1}{2}\mathrm{diag}(w)-\frac{1}{2}\mathrm{diag}(\alpha)\right)$，则相应的神经网络为 $N=(w,\boldsymbol{\theta})$ 对应的能量函数是 $\mathrm{E}(\boldsymbol{u})=\boldsymbol{u}^{\mathrm{T}}w\boldsymbol{u}-2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\theta}$；与 $N=(w^0\oplus w^1,\boldsymbol{\theta})$ 对应的能量函数为 $\mathrm{E}(\boldsymbol{u},\boldsymbol{v})=\boldsymbol{u}^{\mathrm{T}}w^0\boldsymbol{u}+2\boldsymbol{u}^{\mathrm{T}}w^1\boldsymbol{v}-2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\theta}$. 这里 $\boldsymbol{\alpha}=\mathrm{diag}(\alpha_1,\alpha_2,\cdots,\alpha_n)$.

定理 4. 优化问题 $E(\boldsymbol{u})=\boldsymbol{u}^{\mathrm{T}}w\boldsymbol{u}-2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\theta}$，如上分解形式的二阶延迟离散 Hopfield-型网络为 $N=(w^0\oplus w^1,\boldsymbol{\theta})$，如果存在 $\boldsymbol{\alpha}=\mathrm{diag}(\alpha_1,\alpha_2,\cdots,\alpha_n)$，满足

$$w_{ii}-\alpha_i\geqslant\frac{1}{2}\sum_{j=1}^{n}|w_{ij}-w_{ji}|$$

则可以用 GURD1(2) 演化规则进行求解优化问题.

定理的证明只需注意到对 w 的分解 $w=w^0+2w^1$，并应用定理 1 和定理 2 即可. 详细证明略去.

定理的结果拓宽了用神经网络求解优化问题的范围,推广了文献[5]中引理1(C3)的结果(即取 $\boldsymbol{\alpha}=\mathrm{diag}(0,0,\cdots,0)$时).

4.4 稳定态与能量函数的关系

以上的研究表明,网络的一般演化规则使得能量函数单调不减,即当网络达到稳定态时,能量函数在"一定意义上"达到相对极大值点.虽然我们没有给出"一定意义上"的确切定义,但稳定态是能量函数的相对极大值点,这一点从定理1和2得到了证实.但对网络实施一般异步演化规则时,在"一定意义上"达到的能量函数的相对极大值点有确切的定义,即网络的稳定态是能量函数 $E(\boldsymbol{u},\boldsymbol{v})$的 Hamming 距离为 $r=1$ 的极大值点.下面定理将说明这一点.

定理 5. 延迟离散 Hopfield-型网络 $N=(w^0\oplus w^1,\boldsymbol{\theta})$, w^0 是 $n\times n$ 阶对称矩阵, w^1 是 $n\times n$ 阶矩阵且对角元素满足 $w_{ii}^1\geqslant\sum_{\substack{j=1\\j\neq i}}^{n}|w_{ji}^1|$,若网络依一般异步演化规则(对任意的 k, $|L(k)|=1$)进行演化并满足 $\bigcup_{k>k_0}L(k)=\{1,2,\cdots,n\}$,则网络的稳定态是能量函数 $E(\boldsymbol{u},\boldsymbol{v})$的 Hamming 距离为 $r=1$ 的极大值点.

5 结论

本文给出了离散 Hopfield-型网络的一般演化规则,推广了文献[6]的演化规则,使之适用于非对称连接权阵的情况并刻画了对角元素的意义.我们还针对延迟离散 Hopfield-型网络提出了一般演化规则 GURD1 和 GURD2,它的优点在于求解优化问题的分解策略上,实现了延迟项的的调节作用和对非对称权阵的分解转化思想.在新提出的一般演化规则的意义下,给出了延迟离散 Hopfield-型网络和离散 Hopfield-型网络的收敛性定理,更重要的意义是它们在求解优化问题上得到统一.在不同演化规则意义下所给出的收敛性定理在神经元的输出函数的选择方面均有突破,使一直作为收敛性定理必要条件的"连接权阵的对角元非负"[1,2,4,5,7~13]的限定可以去掉.虽然在一般演化规则的意义下能量函数是相对的局部极大值点,对优化问题还是颇有指导意义的;在一般演化规则进行演化时,能量函数是 r=1 距离的正规能量函数,其实验结果是满意的.

References

1 Shrivastava Y. Guaranteed convergence in a class of Hopfield networks. *IEEE Transactions on Neural Networks*, 1992, **3**(6): 951~961

2 Arun J. Approximating maximum clique with a Hopfield network. *IEEE Transactions on Neural Networks*, 1995, **6**(3): 724~735

3 Ramanujam J, Sadayappan P. Mapping combinaterial optimization problems onto neural networks. *Information Science*, 1995, **82**: 239~255

4 Cosnard M, Goles E. Discrete state neural networks and energies. *Neural Networks*, 1997, **10**(2): 327~334

5 Xu Zong-Bin, Hu Guo-Qing, Kwong Chung-Ping. Asymmetric Hopfield-type networks: Theory and applications. *Neural Networks*, 1996, **9**(3): 483~501

6 Sun Yi. A generalized updating rule for modified Hopfield neural network for quadratic optimization. *Neurocompting*, 1998, **19**: 133~143

7 Qiu Shen-Shan, Liu Yong-Qing. Convergence of discrete Hopfield -type neural networks with delay. In: Proceedings of the 2000 American Control Conference, Hyatt Regency Chicago, Illinois, USA: 2000, **1**: 658～659

8 Qiu Shen-Shan, Xu Xiao-Fei, Liu Ming-Zhu *et al*. Convergence of discrete Hopfield-type neural network with time-delay in a serial mode. *Journal of Computer Research & Development*, 1999, **36**(5): 546～552(in Chinese)

9 Qiu Shen-Shan, Tsang E C C, Yeung D S. Stability of discrete Hopfield neural networks with time-delay. In: Proceedings of the 2000 IEEE International Conference on System, Man Cybernetics, 2000. 200～205

10 Vidyasagar M. Discrete optimization using analog neural networks with discontinuous dynamics. *International Series of Numerical Mathematics*, 1996, **121**: 107～112

11 Bruck J, Goodman J W. A generalized convergence theorem for neural networks. *IEEE Transaction on information theory*, 1988, **34**(5): 1089～1092

12 Lee Donq-Liang. New stability conditions for Hopfield networks in partial simultaneous update mode. *IEEE Transactions on Neural Networks*, 1999, **10**(4): 975～978

13 Bruck J. On the convergence properties of the Hopfield model. *Proceedings of IEEE*, 1990, **78**(10): 1579～1585

邱深山　1999 年于哈尔滨工业大学计算机系获工学博士学位，同时在香港理工大学计算机系做合作研究. 1999～2001 年在华南理工大学做博士后研究. 研究领域是神经网络、机器学习、SVM、模糊智能学习理论、软件产业发展研究.

(**QIU Shen-Shan**　Received the Ph. D. degree from Harbin Institute of Technology in 1999, at the same time he was invited as research associate at department of computing, Hong Kong polytechnic University. Form 1999 to 2001, he worked as postdoctoral position in department of Automatic science & engineering of South China University of Technology. His research interests include neural network, machine learning, support vector machines, fuzzy intelligent learning theory, and software industry development research.)

邓飞其　1997 年于华南理工大学获工学博士学位，现为华南理工大学自动控制科学与工程学院教授、博士生导师. 研究领域是系统工程、信息系统分析和设计、神经网络、SVM、复杂系统控制和大型项目的管理咨询及开发.

(**DENG Fei-Qi**　Received the Ph. D. degree in engineering from South China University of Technology (SCUT) in 1997. He is currently a professor and Ph. D. student supervisor of college of Automation science and engineering of SCUT′. His research interests include systems engineering, analysis and design of information systems, neural network, support vector machines, and complex large system control theory.)

刘永清　1955 年毕业于复旦大学数学系，现为华南理工大学自动控制科学与工程学院教授、博士生导师. 研究方向为大系统理论、系统工程.

(**LIU Yong-Qing**　Graduate from Department of mathematics in Fudan University 1955. He is currently a professor and Ph. D. student supervisor of college of Automation science and engineering, SCUT. His research interests include systems engineering and large systems theory.)