43-45 联想记忆网络 动态学习算法 神经网络 (11

多模式对模糊联想记忆网络的动态学习算法*)

A Dynamical Learning Algorithm for Multiple Samples Fuzzy Associative Memory Neural Networks

TP18

肖 平[1] 冯久超[2] 余英林[1]

（华南理工大学电子与通信工程系 广州 510641）[1] （西南师范大学物理系 重庆 400715）[2]

摘 要 In the first, a quick dynamical adjusting learning algorithm for fuzzy associative memory networks is proposed, and a sufficient and necessary condition is given. Then, a dynamical exponent adjusting algorithm is developed, which takes the weights obtanied by quick dynamical algorithm as initial iteration value. The simulations demonstrate the advantage of the proposed algorithm.

关键词 Neural netwrok, Fuzzy associative memory, Learning algorithm, Dynamical adjusting

Kosko 在结合模糊系统与神经网络的基础上提出的最大最小联想记忆网络[1]是二层前馈网络，为了提高其记忆性能，已经提出了一些改进的算法。其中主要有正交编码算法[2]，最大连接权算法[3]，梯度下降算法[4~6]。梯度算法需要较复杂的迭代运算，设计出的网络不具有“学习”能力，即当网络需要存储一个新的训练模式时，就必须重新设计整个网络权矩阵。最大连接权法在网络存在精确连接权的条件下，对存储模式因信息丢失的联想效果最好，但对随机噪声畸变模式其联想效果却不甚理想。尤其在网络不存在精确解的条件下，网络的记忆性能显得较差。文[2]的正交编码算法是一种能有效地结合权值分析解与数值解的算法。这种算法的记忆容量要好于最大权算法。

本文首先给出联想记忆网络的一种快速动态调整算法，而后发展了高容量记忆网络的一种动态指数调整学习算法，它保持了正交编码算法快速简便等特点。通过参数的动态调整，可以提高网络的记忆容量，而且对于输入不同类型的畸变存储模式，连接权采用动态调整，可使其联想效果达到最好。

一、最大最小双向联想记忆模型及其相关分析

Kosko 给出的最大最小模糊联想记忆网络模型为：

$$A_k \circ W = B_k, k=1,2,\cdots,p$$

用逐点表示，即为：

$$\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})=b_{kj}, j=1,2,\cdots,m, k=1,2,\cdots,p \qquad (1)$$

其中 $A_k=(a_{k1},a_{k2},\cdots,a_{kn})$，$B_k=(b_{k1},b_{k2},\cdots,b_{km})$分别为输入和输出层的模糊向量；$W=(w_{ij})_{n*m}$为网络的连接权；n 和 m 分别为输入和输出层神经元个数；p 为训练模式对数；“$\vee$”和“$\wedge$”表示最大和最小运算。

1.1 快速动态调整学习算法及模糊双向联想记忆模型

设 $W_k=(w_{ij}^k)_{n*m}$和 $W=(w_{ij})_{n*m}$分别是第 k 个模式对和整个联想记忆模型(1)的关系矩阵。定义：

$$w_{ij}^k(\lambda)=a_{ki}\theta b_{kj}=\begin{cases}1, & \text{if } a_{ki}=b_{kj} \text{ or } b_{kj}-a_{ki}\geqslant\lambda \\ a_{ki}\wedge b_{kj}, & \text{else}\end{cases}$$

$$w_{ij}(\lambda)=\bigwedge_{k=1}^{p}(w_{ij}^k(\lambda))=\bigwedge_{k=1}^{p}(a_{ki}\theta b_{kj}) \qquad (2)$$

其中 λ 是一个动态参数，很显然，当 λ=0 时，上述学习算法即为文[3]的最大权编码算法。

最大最小算子突出的是主因素作用，对于输入不完整即丢失部分信息的记忆模式，网络能较好地联想出其相应的模式；而对于模式分量值变大时就较难正确联想。因此，在设计网络时，主要应考虑如何使网络对这种畸变模式容错性较强。以单个模式对为例，当采用最大权编码，若 $a_{ki}<b_{kj}$，则 $w_{ij}^k=1$。从而 $a_{ki}\wedge w_{ij}^k<b_{kj}$，当 a_{ki}变大至 $a_{ki}>b_{kj}$时，则 $\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij}^k)\geqslant a_{ki}\wedge w_{ij}^k>b_{kj}$，而取 λ>1，由式(2)，若 $a_{ki}<b_{kj}$，则 $w_{ij}^k=a_{ki}\wedge b_{kj}=a_{ki}$。因此不论 a_{ki}如何变大，都有 $\bigvee_{l=1,l\neq i}^{n}(a_{kl}\wedge w_{lj}^k)\vee(a_{ki}\wedge w_{ij}^k)=\bigvee_{l=1,l\neq i}^{n}(a_{kl}\wedge w_{lj}^k)\vee a_{ki}=\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij}^k)=b_{kj}$。

若模型(1)的连接权矩阵 W 由(2)式确定，则前

*)本项目受国家攀登计划和国家自然科学基金资助。肖平 博士生，副教授。冯久超 副教授。余英林 教授，博士生导师。

馈模糊联想记忆模型(1)可以发展成为一种模糊双向联想记忆模型，即：

$$\begin{cases} A_k \circ W = B_k' \\ B_k \circ R = A_k' \end{cases} \quad k=1,2,\cdots,p。$$

逐点表示就为：

$$\begin{cases} \bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij}) = b_{kj}',\ j=1,2,\cdots,m,\ k=1,2,\cdots,p \\ \bigvee_{j=1}^{m}(b_{kj}\wedge r_{ji}) = a_{ki},\ i=1,2,\cdots,n,\ k=1,2,\cdots,p \end{cases} \tag{3}$$

上式中 $W=(w_{ij})_{n\times m}$，$R=(r_{ji})_{m\times n}$，$A_k'=(a_{k1}',a_{k2}',\cdots,a_{kn}')$ 和 $B_k'=(b_{k1}',b_{k2}',\cdots,b_{km}')$ 分别是神经元 A_k 和 B_k 的下一个状态向量。很显然，由式(3)可知，一个模糊模式向量 (A_k,B_k) 是双向联想记忆模型(3)的一个稳定吸引子当且仅当 $\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})=b_{kj}$ 和 $\bigvee_{j=1}^{m}(b_{kj}\wedge r_{ji})=a_{ki}$，$k\in\{1,2,\cdots,p\}$ 同时成立。

1.2 双向记忆模型的稳定性条件及容错性分析

定义 1[8] 反馈网络中一个状态序列 $\{p_1,p_2,\cdots,p_s\}$ 称为极限环，如果对网络的状态转移因子 F，有 $F(p_1)=p_2$，$F(p_2)=p_3\cdots$，$F(p_k)=p_{k+1}$，$\cdots$，$F(p_s)=p_1$，并且其中不含有具有这种性质的子序列，s 称为极限环的长度。

定理 1 模糊双向联想记忆网络的任一初始状态矢量，最终都将收敛到某一极限环。

证明方法与文[8]类似，从略。

定理 2 设 (A_k,B_k) 是 p 个向量模式对，式中 $A_k=(a_{k1},a_{k2},\cdots,a_{kn})\in[0,1]^n$，$B_k=(b_{k1},b_{k2},\cdots,b_{km})\in[0,1]^m$，则对于任意给定的 $k\in\{1,2,\cdots,p\}$，模式向量 B_k 可由向量 A_k 一次联想出来的充要条件是对于任意的 $j\in\{1,2,\cdots,m\}$，都至少存在一个 $i\in\{1,2,\cdots,n\}$ 使得：

$$b_{kj}\leqslant a_{ki}\wedge w_{ij} \tag{4}$$

其中 w_{ij} 是由学习规则(2)决定。

证明：(必要性)设 $A_k\circ W=B_k$，即 $\bigvee_{k=1}^{n}(a_{ki}\wedge w_{ij})=b_{kj}(j=1,2,\cdots,m)$。如果存在一个 $j\in\{1,2,\cdots,m\}$，使得对于所有的 $i\in\{1,2,\cdots,n\}$，都有 $b_{kj}>a_{kj}\wedge w_{ij}$，那么 $b_{kj}>\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})=b_{kj}$。矛盾。

(充分性)假定对于任给的 $j\in\{1,2,\cdots,m\}$，都至少存在一个 $i\in\{1,2,\cdots,n\}$，使得 $b_{kj}\leqslant a_{ki}\wedge w_{ij}$，则可得 $b_{kj}\leqslant a_{ki}$ 和 $b_{kj}\leqslant w_{ij}$，下面分两种情况讨论之。

(i) $b_{kj}<a_{ki}$ 和 $b_{kj}\leqslant w_{ij}$。则由规则(2)，$b_{kj}\leqslant w_{ij}=\bigwedge_{k=1}^{p}(a_{ki}\theta b_{kj})\leqslant a_{ki}\theta b_{kj}=b_{kj}$，因此 $w_{ij}=b_{kj}$。故 $\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})=\bigvee_{i=1}^{n}(a_{ki}\wedge b_{kj})=\bigvee_{i=1}^{n}b_{kj}=b_{kj}$。(ii) $b_{kj}=a_{ki}$ 和 $b_{kj}\leqslant w_{ij}$。则 $\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})\geqslant a_{ki}\wedge w_{ij}\geqslant a_{ki}\wedge b_{kj}=b_{kj}$。又 $\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})\leqslant\bigvee_{i=1}^{n}(a_{ki}\wedge 1)\leqslant\bigvee_{i=1}^{n}(b_{kj}\wedge 1)=b_{kj}$。因此 $\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})=b_{kj}$。综合(i)(ii)，可得 $\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})=b_{kj}$。 □

定理 3 设 (A_k,B_k) 是 p 个训练模式对，则对于每个 $k\in\{1,2,\cdots,p\}$，模式对 (A_k,B_k) 在一次循环迭代后成为双向联想记忆模型(3)的一个稳定态的充要条件是同时满足条件(a)和(b)：

(a)对于任意给定的一个 $j\in\{1,2,\cdots,m\}$，至少存在一个 $i\in\{1,2,\cdots,n\}$，使得 $b_{kj}\leqslant a_{ki}\wedge w_{ij}$

(b)对于任给的一个 $i\in\{1,2,\cdots,n\}$，至少存在一个 $j\in\{1,2,\cdots,m\}$，使得 $a_{ki}\leqslant b_{kj}\wedge r_{ji}$ (5)

证明方法与定理 1 类似，从略。

由上式(4)，(5)可知，对于自联想模式：$a_{ki}=b_{ki}$，显然有 $a_{ki}\leqslant a_{ki}\wedge w_{ii}$，因此不论 λ 为何值，快速动态调整算法均能完整记忆所有训练样本，而不论训练样本的数目及强弱如何。

推论 1 设某个模式对 (A_k,B_k)，$k\in\{1,2\cdots,p\}$，是双向联想记忆模型(3)的一个稳定吸引子，$A=(a_1,a_2,\cdots,a_n)$ 是模式向量 A_k 的一个畸变模式。如果对于任意一个 $j\in\{1,2,\cdots,m\}$，存在一个 $i\in\{1,2,\cdots,n\}$，使得 $b_{kj}=a_i\wedge w_{ij}$ 且 $a_l\leqslant a_{kl}(1\neq i)$，则模式向量 B_k 可由畸变模式 A 迭代一次联想出来。

证明：首先证明模式向量 B_k 的第 j 个分量 b_{kj} 可以由向量 A 迭代一次回想出来。由于 $\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})=b_{kj}$，且 $a_l\leqslant a_{kl}(l\neq i)$，则有 $a_l\wedge w_{lj}\leqslant a_{kl}\wedge w_{lj}(l\neq i)$。因此 $\bigvee_{l=1,l\neq i}^{n}(a_l\wedge w_{lj})\leqslant\bigvee_{l=1,l\neq i}^{n}(a_{kl}\wedge w_{lj})\leqslant b_{kj}$，而由假设 $b_{kj}=a_i\wedge w_{ij}$，则 $b_{kj}=a_i\wedge w_{ij}\leqslant\bigvee_{i=1}^{n}(a_i\wedge w_{ij})\leqslant\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})=b_{kj}$。所以 $\bigvee_{i=1}^{n}(a_i\wedge w_{ij})=b_{kj}$。由于 j 的任意性，故结论成立。

从推论 1 可以看出，只要保留某个记忆模式向量的一些主要信息，即使丢失了其它部分信息，则权值由学习规则(2)所确定的双向联想记忆网络仍可以完整地回想起其相应的模式向量。这正体现了 max 和 min 合成的模糊神经元突出的是主因素作用。

二、动态指数调整学习算法

由于在实际问题中，训练模式集的获取受人为因素的影响，并不一定满足条件(4)或(5)，因此，倘若以原模式的相似模式代替原模式进行编码，而能使整个系统更加相容，这样就确定了系统连接权的一个近似解，最简单的方法就是以训练模式对(A，B)的语言变量模式 (A^α,B^α)（在语言值模糊逻辑中，α 称为语言变量，如很，一般，相当等）代替(A，B)进行编码。需要说明的是文[2]的正交编码的思想即源于此，但该文在阐述方法的合理性上不够妥当。本文网

络的连接权采用规则(2)确定，即 $w_{ij}=\bigwedge_{k=1}^{p}(a_{ki}^{\alpha_k}\theta b_{kj}^{\alpha_k})$。为了确定 α_k，我们发展了一种称为动态指数调整学习算法。算法总结如下：设全局误差函数 E：

$$E=\frac{1}{p}\sum_{k=1}^{p}\sum_{i=1}^{m}(\bar{b}_{kj}-b_{kj})^2,\ \bar{b}_{kj}=\bigvee_{i=1}^{n}(a_{ki}\wedge w_{ij})。$$

(1)选取动态调整参数 $\lambda=0$。

(2)给定一组指数初始值 $\alpha_k(t)=1(k=1,2,\cdots,p$，p 为训练样本数)，迭代次数 M 及迭代精度 ε。

(3)产生高斯随机变量 $\xi_k(t)$，若 $\alpha'_k(t)=\alpha_k(t)+\mu\cdot\xi_k(t)\in X$。($\mu$ 是高斯方差，X 是搜索区域，t 为迭代次数)，则转(4)；否则 $\alpha'_k(t)=\alpha_k(t)$。

(4)计算 $w_{ij}(t)=\bigwedge_{k=1}^{p}(a_{ki}^{\alpha'_k(t)}\theta b_{kj}^{\alpha'_k(t)})=,i=1,2,\cdots,m$。

(5)若 $E(t)<E(t-1)$，则 $\alpha_k(t+1)=\alpha'_k(t)$，否则 $\alpha_{k+1}(t+1)=\alpha_k(t)$。

(6)若 $t=M$ 或 $E(t)<\varepsilon$，则结束；否则 $t=t+1$，转(2)。

(7)调整 $\lambda=\lambda+\beta(\beta>0$ 为一调整量)，重复(1)至(6)。

最后选取最适宜的参数 λ。当 $\lambda=0$ 时，动态调整算法即为文[2]的正交编码学习算法。很显然，上述方法与某些神经网络迭代算法不同，网络训练与输入次序无关。

三、模拟结果

实验一 给出了 6 组模糊规则集(a)～(f)。表中记号"√"表示输入输出对应关系，其中 S=小，BS=偏小，M=中，BL=偏大，L=大。记 A=B={小，偏小，中，偏大，大}，例如(a)的意思是"若 a 小，则 b 小；若 a 偏小，则 b 偏大；若 a 中则 b 中；若 a 偏大；则 b 偏大；若 a 大，则 b 大"，论域 X=Y={0,0.5,1,1.5,2,…,12}。用高斯函数定义它们的隶属函数：

$$\mu_S(x)=\begin{cases}1 & x\leqslant 2\\ e^{-(x-2)^2/4} & x>2\end{cases}$$

$$\mu_{BS}(x)=\begin{cases}1 & x\leqslant 4\\ e^{-(x-4)^2/4} & x>4\end{cases}$$

$$\mu_M(x)=e^{-(x-6)^2/4},$$

$$\mu_{BL}(x)=\begin{cases}1 & x\geqslant 8\\ e^{-(x-8)^2/4} & x<8\end{cases}$$

$$\mu_L(x)=\begin{cases}1 & x\geqslant 10\\ e^{-(x-10)^2/4} & x<10\end{cases}$$

表 1 给出了最大权，正交编码及动态指数调整三种学习算法所产生的误差 $E1=E/25$，($n=15,m=25,p=5$)，实验时，迭代次数均为 100 次，从表 1 可看出，最大权算法在模式对不满足条件(4)，(5)时，误差较大，本文的动态指数调整算法优于正交编码算法。

Y/X	S	BS	M	BL	L
S	√				
BS		√			
M			√		
BL				√	
L					√

(a)

Y/X	S	BS	M	BL	L
S				√	
BS	√				
M			√		
BL		√			
L					√

(b)

Y/X	S	BS	M	BL	L
S				√	
BS	√	√			
M			√		
BL					
L					√

(c)

Y/X	S	BS	M	BL	L
S					√
BS					
M			√		
BL	√				
L					

(d)

Y/X	S	BS	M	BL	L
S					√
BS			√	√	
M			√		
BL		√	√		
L	√				

(e)

Y/X	S	BS	M	BL	L
S					√
BS				√	
M		√	√	√	
BL		√			
L	√				

(f)

表 1

规则集	最大权[3]	正交编码[2]	动态调整	
	E1	E1	E1	λ
a	0	0	0	0
b	1.52	1.05	0.91	0.9
c	0.64	0.64	0.54	0.6
d	0	0	0	0
e	1.74	0.81	0.72	1.0
f	1.29	0.60	0.57	0.4

参考文献

[1] B. Kosko, Neural networks and Fuzzy systems: A dynamical approach to machine Intelligence. Prentice Hall. 1991

[2] Fu-lai chung and Tong Lee, Toward a high capacity fuzzy associative memory model, IEEE int. conf. on neural networks. 1994

[3] 范俊波、靳蕃等，模糊联想记忆的一种有效学习算法，电子学报，24(1)1996

[4] 何奉道，多模式对模糊联想记忆学习的优化，计算机科学，21(3)1994

[5] A. Blanco et al., Identification of fuzzy relational equations by fuzzy neural networks, Fuzzy sets and systems, 1995, 71

[6] N. Ikoma et al., Estimation of fuzzy relational matrix by using probabilistic descent method, Fuzzy sets and systems, 1993, 57

[7] Nori Baba, A new approach for finding the global minimum of error function of neural network, Neural network, 1989, 2

[8] 刘增良，刘有才，因素神经网络理论及应用，贵州科技出版社，1994