

# 模糊联想记忆神经网络模型在地震预报中的应用<sup>\*</sup>

王 炜<sup>1)\*</sup> 吴耿锋<sup>2)</sup> 黄冰树<sup>1)\*\*\*</sup> 庄昆元<sup>3)</sup>  
周佩玲<sup>2)</sup> 蒋春曦<sup>1)</sup> 李东升<sup>1)</sup> 周云好<sup>1)</sup>

1) 中国合肥 230031 安徽省地震局

2) 中国合肥 230026 中国科学技术大学

3) 中国上海 200062 上海市地震局

**摘要** 介绍了模糊联想记忆 FAM (Fuzzy Associative Memory) 神经网络模型、FAM 自适应学习算法以及 FAM 推理机的原理，并成功地将其用于“新一代的地震预报专家系统 NGESEP”，使得系统既具有良好的学习功能，又避免了通常神经网络学习知识隐含在权值中不易被人们理解和专家系统解释的缺点。

**主题词** 专家系统 模糊神经网络 模糊联想记忆 积空间聚类

## 引言

专家系统技术正处于由第1代向第2代发展的阶段。目前专家系统技术在地震预报中已有了初步运用，但是于80年代末期研制成功的3个地震预报专家系统都属于第1代系统(庄昆元等, 1991)。第1代专家系统通常由知识库、推理机和解释器三部分构成。第1代专家系统的“瓶颈”问题是知识获取。知识来自于符号表示的手工经验获取，而经验来自于专家的主观认识。知识获取费时、效率低。因此，专家系统的发展依赖于学习和推理技术的进步。目前正在兴起的第2代专家系统为解决知识获取的“瓶颈”，则力图从事实数据中自动提取知识，即实现机器学习。从而使得学习器、知识库、推理机和解释器成为第2代专家系统的必备主件。

80年代中期以来，人工神经网络理论和应用开始重新兴起。由于神经网络较强的学习功能，它与专家系统在技术上具有定量分析与定性分析、自动获取知识与手工获取知识的互补性，因此集两者的学习和推理功能于一体的神经网络专家系统成为第2代专家系统的主流。但是神经网络中的知识通常隐含在权值中，我们很难从权值知道网络编码的内容是什么，因此神经网络不能直接处理结构化知识，即由专家给出的“规则”。这样，专家系统

\* 地震科学联合基金资助课题。

1995-12-14 收到初稿, 1996-11-05 收到修改稿并决定采用。

\*\* 现在地址：上海市地震局。

\*\*\* 现在地址：上海材料研究所。

对推理结果进行解释就十分困难。而模糊系统则是通过模糊语言变量及隶属度描述事件, 使用语言方式来描述命题, 并通过模糊规则和推理由前提得到结论。对于每个输入的响应, 其响应的过程是对每条规则并行地“激活”, 从而得到不同程度满足这些规则的响应, 其过程是一目了然的。这表明模糊系统可以很方便地处理结构化知识, 因此, 专家系统对这类知识的推理过程解释就十分方便。但是模糊系统不具备自学习功能, 通常规则是由专家以语言方式表达出来, 专家对问题认识的深度和综合能力直接影响到模糊系统工作性能的好坏; 另外, 有时即便是经验丰富的专家也很难归纳出一些比较明确而简化的规则。于是近几年来科学家开始将神经网络、模糊系统与专家系统结合起来, 充分发挥各自的优点, 利用神经网络学习功能产生模糊规则, 而利用模糊规则推理方便、过程清晰的特点使专家系统的推理及解释功能较易实现。这是目前在专家系统和神经网络领域的前沿研究课题。在新一代的地震预报专家系统 NGESEP(New generation expert system for earthquake prediction)中使用了模糊联想记忆 FAM 模型, 这使得系统除具有第 1 代专家系统的较强符号推理和解释功能外, 还具有较强的学习功能。

## 1 模糊联想记忆(FAM)模型

### 1.1 模糊联想记忆(FAM)模型

FAM 模型(Kosko, 1991)是一个两层前馈异联想模糊分类器, 或者说是一个模糊映射系统。如图 1 所示, 输入为模糊变量  $\mathbf{A}$ , 它是一个  $n$  维矢量, 通过  $m$  条规则, 映射到  $p$  维模糊矢量  $\mathbf{B}$  上。FAM 系统通过学习可存储任意模糊空间模式对  $(A_k, B_k)$ , 用模糊集表达的第  $k$  个模式对为

$$\begin{aligned} A_k &= \{a_1^k, \dots, a_n^k\} & B_k &= \{b_1^k, \dots, b_p^k\} \\ F_k &= \mathbf{A}^T o \mathbf{B} = \min\{\mu_A(a_i^k), \mu_B(b_j^k)\} \end{aligned} \quad (1)$$

这里,  $\mu_A(a_i^k)$  是模糊变量  $\mathbf{A}$  中第  $i$  个元素的隶属度,  $\mu_B(b_i^k)$  是模糊变量  $\mathbf{B}$  中第  $i$  个元素的隶属度,  $o$  表示合成运算,  $F_k$  是一个  $n \times p$  的模糊矩阵。在 FAM 中,  $(A_k, B_k)$  代表了第  $k$  条推理规则, 它是蕴涵句 “If  $X$  is  $A_k$  then  $Y$  is  $B_k$ ” 的简写形式。FAM 规则也可表达多前提规则, 如  $(A_k, B_k; C_k)$ , 即表示 “If  $X$  is  $A_k$  and  $Y$  is  $B_k$  then  $Z$  is  $C_k$ ”。显然, 这对于表达地震预报中的多前提(如根据空区长轴和持续时间预报未来地震震级)是非常适合的。

由上述可知, 每一条规则都对应了一个模糊矩阵  $F_k$ , 各条规则独立存放在 FAM 系统中, 这样做的好处是可以灵活方便地增、删或修改规则, 在推理时又可以清楚地了解 FAM 系统中每条规则对输出模糊矢量  $\mathbf{B}$  的贡献是多少。

### 1.2 FAM 推理机

由 FAM 学习网络得到的规则可供 FAM 推理机进行推理。如图 1 所示, 若在 FAM 系统中有  $m$  条规则, 输入为一个  $P$  维的模糊矢量  $\mathbf{A}$ (通常输入为确切的数值, 需经模糊化处理)。它不同程度地并行激活相应的 FAM 规则  $F_i$ , 并得到输出  $B'_i$ 。这样,  $m$  条规

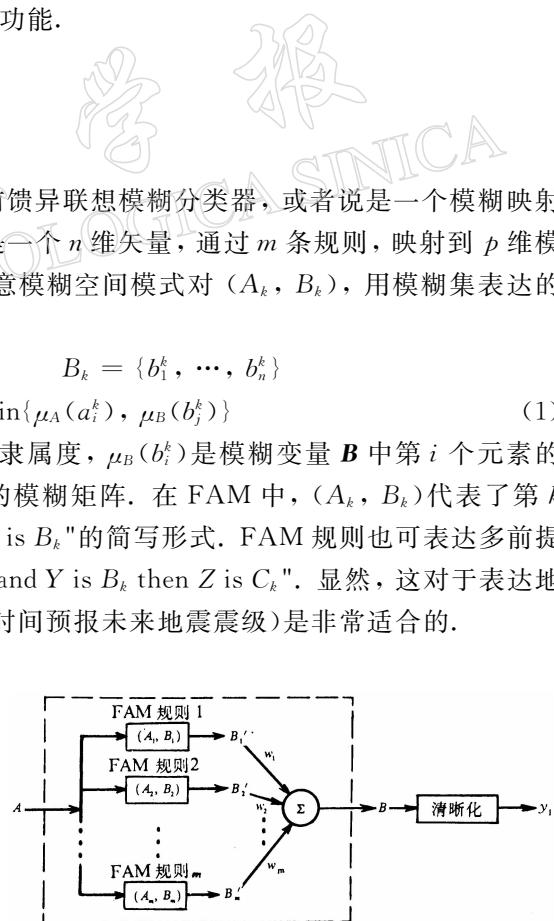


图 1 模糊联想记忆 FAM 系统

则可能产生  $m$  个子结论  $B'_1, B'_2 \dots, B'_m$ . 将  $m$  个子结论按下式进行归一化加权求和便得到 FAM 的最后输出  $\mathbf{B}$

$$\mathbf{B} = W_1 B'_1 + W_2 B'_2 + \dots + W_m B'_m \quad (2)$$

其中, 权系数  $W_k$  反映了第  $k$  条规则在推理或联想中的强度值. 最后经过去模糊化得到具体的数值  $B^0$ . 通常去模糊化(清晰化)方法一般有两种:

(1) 最大隶属函数法. 若以  $\mu_B$  表示  $B$  的隶属函数, 则按峰值法确定的数值  $B^0$  为

$$B^0 = y_k \quad (3)$$

其中,  $y_k$  满足  $\mu_B(y_k) = \max \{\mu_B(y_j)\}, 1 \leq j \leq p$

最大隶属函数法存在两个主要问题: 第一, 当  $\mu_B$  的形状有多个等高峰值时, 按上式求得的  $y_k$  不唯一; 第二, 该方法在很大程度上忽略了  $\mu_B$  形状所包含的信息.

(2) 重心法. 与最大隶属函数法相比较, 重心法注重的不是  $\mu_B$  的峰值, 而是  $\mu_B$  的整个形状. 这时  $B^0$  为

$$B^0 = \sum_{j=1}^p y_j \mu_B(y_j) / \sum_{j=1}^p \mu_B(y_j) \quad (4)$$

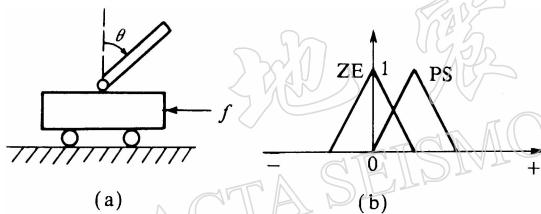


图 2 倒立摆(a)及  $\theta$  的语言值 ZE 和 PS 的隶属函数(b)

本系统使用的去模糊方法为重心法. 下面将通过倒立摆的平衡控制问题来说明上述推理过程. 如图 2a 所示, 倒立摆是指一支铰接在一个可移动小车上的摆, 我们的目的是希望通过不断地对小车施以向左或向右的作用力, 以使摆能够从不同的初始状态回到并保持在竖直的平衡位置. 对于这样一个控制问题, 传统方法是先建立起整个系统的数学模型, 然后求出外力作用的数学解. 现在我们用模糊控制的方法

来解决这一问题. 选择偏离角  $\theta$  和  $\theta$  的变化速度(角速度)  $\Delta\theta$  为观测变量, 摆的位置在垂线右方时,  $\theta$  取正值; 反之,  $\theta$  取负值. 外加作用力  $f$  为控制变量,  $f$  的正方向向左, 如图 2a 中箭头所示. 当摆向垂线右方偏落时,  $\Delta\theta$  为正值, 这时我们希望  $f$  为负值; 当摆从垂线左方偏落时,  $\Delta\theta$  为负值, 这时我们希望  $f$  为正值. 从这里我们可以看到, 当摆偏落时,  $f$  的本质作用是抵消  $\Delta\theta$ .

如果视  $\theta$ ,  $\Delta\theta$  和  $f$  为模糊变量, 则它们的取值将是一些语言值, 这些语言值对应了论域  $R$  上的一些模糊集合. 例如对于模糊变量  $\theta$ , 规定它的全部语言值为: 负大、负中、负小、零、正小、正中、正大, 或分别表示为: NB, NM, NS, ZE, PS, PM, PB. 这些符号的意义如下:

(负大)NB: Negative Big (负中)NM: Negative Medium (负小)NS: Negative Small  
(零)ZE: Zero

(正小)PS: Positive Small (正中)PM: Positive Medium (正大)PB: Positive Big

NB, NM, …, PB 这 7 个语言的语义对应了论域  $R$  上的 7 个模糊集合. 例如, ZE 和 PS 的隶属函数可由图 2b 示意.

对于模糊变量  $\Delta\theta$ , 也可规定它的全体语言值为 NB, NM, NS, ZE, PS, PM, PB. 但

是,  $\theta$  的语言值 NB 和  $\Delta\theta$  的语言值 NB 是两个不同的概念, 它们对应了  $R$  上两个不同的模糊集合, 如此等等. 最后, 对于模糊变量  $f$ , 这里也规定它取上述 7 个语言值.

在倒立摆的平衡控制问题中, 我们的目的是力图使  $\theta$  和  $\Delta\theta$  同时保持取零值. 例如, 当摆位于垂线上( $\theta=0$ )且在向左偏落( $\Delta\theta<0$ )时, 应施以正向的作用力( $f>0$ ); 摆向左偏落的速度越快, 作用力  $f$  就应该越大. 凭借这类常识性经验, 很容易对倒立摆的平衡控制问题给出若干条控制规则, 例如(ZE, NS; PS), (ZE, NM; PM), (ZE, ZE; ZE)等等. (ZE, NS; PS)代表了如下的蕴涵句:

$$\text{如果 } \theta = \text{ZE} \text{ 且 } \Delta\theta = \text{NS} \quad \text{则 } f = \text{PS}$$

这里, 我们一共给出了 15 条控制规则, 这

15 条控制规则就构成了一套控制策略.

假定我们得到了一组观测值  $\theta=15$ ,  $\Delta\theta=-10$ , 那么系统将如何工作并最后得到控制作用  $f$  的值呢? 首先看看输入观测数据(15, -10)是如何作用于 FAM 中的规则(ZE, ZE; ZE)的. 如图 3,  $\mu_{\text{ZE}}^{\theta}(15)=0.2$ ,  $\mu_{\text{ZE}}^{\Delta\theta}(-10)=0.5$ , (15, -10)作用于规则(ZE, ZE; ZE)后得到的模糊集合为

$$\min[\mu_{\text{ZE}}^{\theta}(15), \mu_{\text{ZE}}^{\Delta\theta}(-10)] \wedge \mu_{\text{ZE}}^f(f)$$

$$= 0.2 \wedge \mu_{\text{ZE}}^f(f)$$

该模糊集合由图中的阴影区表示. 图 3 还显示了(15, -10)是如何作用于规则(PS, ZE; NS)的, (15, -10)作用于(PS, ZE; NS)后得到的模糊集合为

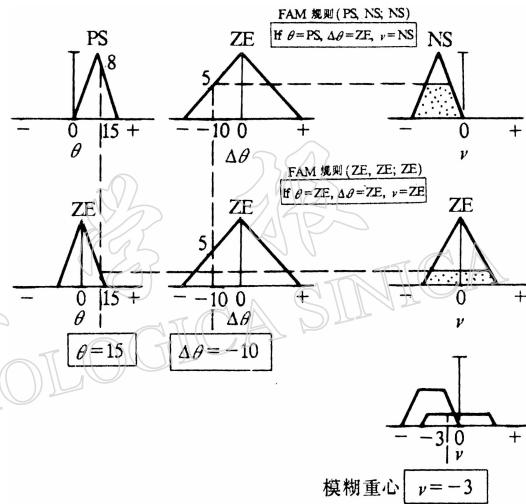


图 3 FAM 重心法计算过程

$$\min[\mu_{\text{PS}}^{\theta}(15), \mu_{\text{ZE}}^{\Delta\theta}(-10)] \wedge \mu_{\text{NS}}^f(f) =$$

$$\min(0.8, 0.5) \wedge \mu_{\text{NS}}^f(f) = 0.5 \wedge \mu_{\text{NS}}^f(f)$$

类似地, (15, -10)还同时作用其余 13 条规则, 最后共得到 15 个模糊集合. 对这 15 个模糊集合进行归一化加权求和, 得到 FAM 的输出模糊集合. 最后, 利用重心法对输出模糊集合进行去模糊处理, 便得到控制器的输出控制量  $f$ , 它为论域  $R$  上的一个确切值.

## 2 FAM 神经网络模型在地震预报中的应用

### 2.1 地震预报规则的产生

以上介绍了 FAM 模型和推理机, 下面将通过实例进一步介绍 FAM 神经网络模型与学习算法. 在地震预报中, 对于各单项预报指标, 只要在有充分样本情况下, 都可使用积空间聚类方法学习得到相应的知识. 具体实现时一般采用自适应二进制输入输出 BIOFAM (Binary Input-Output FAM)方法(Kosko, 1991). 由于输入/输出量通常均为非模糊的确切数值, 这些确切数值所对应的模糊矢量实际上是一些二进制矢量(布尔矢量), 所以相应地称 FAM 系统为 BIOFAM. 这时, 当前提和结论仅为一个模糊集时, 积空间或突触向量为二维. 为说明问题方便起见, 以下以孕震空区长轴  $L(\text{km})$  与未来地震震级  $M$  的关系为例,

说明预报规则的产生过程：

(1) 确定输入和输出(前提与结论)语言值集合. 设  $X$  为输入论域(空区长轴), 根据实际震例资料情况规定输入语言值集合为  $\{L_1, L_2, L_3\}$ ,  $L_1$  表示较短,  $L_2$  表示中等,  $L_3$  表示较长; 设  $Y$  为输出论域(地震强度), 规定输出值集合为  $\{M_1, M_2, M_3\}$ ,  $M_1$  为 5 级左右地震,  $M_2$  为 6 级左右地震,  $M_3$  为 7 级左右地震. 它们的隶属函数为如图 4(a)所示的函数. 这样, 就构成了如图 4(b)所示的  $X \times Y$  积空间, 每个单元对应了  $X$  和  $Y$  上的两个模糊集合, 表示一条虚规则. 由图可知有 9 条虚规则.

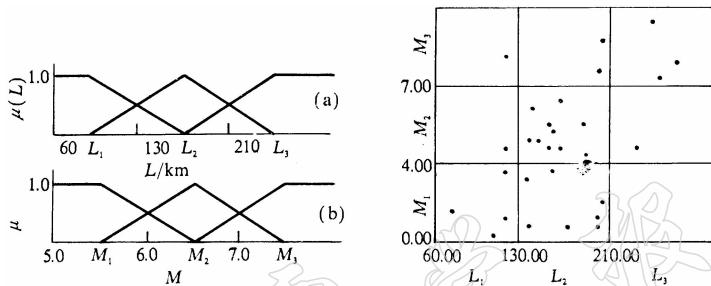


图 4 孕震空区长轴  $L$  与震级  $M$  模糊子集的隶属函数(a)及学习样本在积空间分布(b)

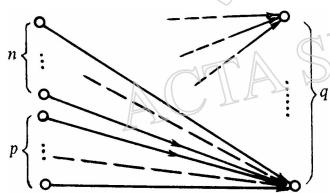


图 5 用于规则形成的单层前向网络

(2) 建立一个单层前向神经网络, 该网络含 2 个输入神经元和  $q$  个输出神经元(图 5). 一般取  $q$  大于样本数及虚规则数, 本例中  $q=49$ . 二维的竞争突触向量  $M_1, M_2, \dots, M_{49}$ , 对应于平面  $X \times Y$  上的 49 个点.

(3) 用竞争自适应向量量化 AVQ(Adaptive Vector-Quantization)方法对输入样本  $(X_i, Y_i)$ , 进行训练(本例中有 40 个样本对). 经过聚类, 根据 49 个突触向量在  $X \times Y$  空间上的分布(图 6a)评估了相应的 FAM 规则, 即聚类定义了规则.

(4) 计算虚规则的权值  $W_{ij} = K_{ij}/q$ ,  $K_{ij}$  为落入某  $F_{ij}$  单元的突触向量数,  $q$  为突触向量总数.

(5) 选择权值大于某阈值的规则为最后确定的 FAM 规则. 本系统将权值  $W_{ij} < R \cdot \sum_{k=1}^n W_k/n$  的虚规则删除, 并将其权值均匀分配给保留的虚规则权值中. 这里  $W_k$  为权值, 其数量为  $n$ ,  $n \leq q$ ;  $R$  为系数, 本系统一般取  $R=0.5$  左右.

经过上述学习过程可以得到一条“主规则”及如下 4 条 FAM 规则(图 6b):

**规则 1** 如果孕震空区长轴较短( $L_1$ ), 那么未来可能会发生 5 级左右地震( $M_1$ ), 权值 0.14.

**规则 2** 如果孕震空区长轴长度中等( $L_2$ ), 那么未来可能会发生 5 级左右地震( $M_1$ ), 权值 0.33.

**规则 3** 如果孕震空区长轴长度中等( $L_2$ ), 那么未来可能会发生 6 级左右地震( $M_2$ ), 权值 0.41.

**规则 4** 如果孕震空区长轴较长( $L_3$ ), 那么未来可能会发生 7 级左右地震( $M_3$ ), 权值 0.12.

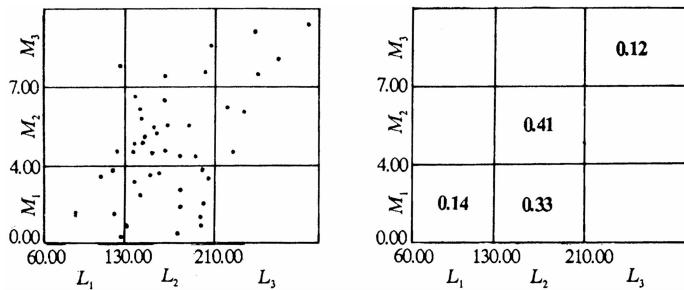


图 6 学习完成后的突触矢量分布(a)及产生的规则(b)

作为验证, 我们根据 40 个地震样本得到一空区长轴  $L$  与主震震级  $M$  的经验关系(由线性最小二乘法得到):  $M = 4.24 \lg L - 3.19$ . 若已知实际孕震空区长轴  $L$ , 则由上述经验公式和学习得到的 FAM 规则推理得到的结果如表 1. 可见, 两种方法得到的结果相近.

表 1 结果比较

长轴 $L/\text{km}$	100	130	160	200	280
震级 $M$ (经验公式)	5.3	5.8	6.2	6.6	7.2
震级 $M$ (FAM 推理)	5.7	5.9	6.1	6.2	7.5

## 2.2 竞争 AVQ 学习算法

在积空间聚类过程中, 我们采用非监督竞争 AVQ 算法. 由于地震学习样本通常较少, 所以一般不要取部分样本值作为突触矢量的初始值; 另外, 为克服学习结果与突触矢量初始值有关的问题, 在学习过程中采用了 Kohonen 自组织映射(Kohonen, 1984)算法中的思想, 学习时不仅对竞争获胜突触向量进行修改, 而且对其邻域中的所有突触向量进行修改, 并且邻域半径随学习次数增大而减小. 该方法具体过程如下:

(1) 突触矢量初始化. 产生  $q$  组任意小的随机数为初始突触矢量  $M_{ij}$ , 并设定邻域为整个输出平面.

(2) 提供学习样本  $X_i(t)$ ,  $0 \leq i \leq N$ ,  $N$  为学习样本数.

(3) 计算输入样本到各输出节点  $j$  之间的距离  $d_j$ .

$$d_j = \| X_i(t) - M_{ij}(t) \| \quad (5)$$

最小的  $d_j$  对应的突触向量为获胜向量, 选择节点  $j^*$  作为最小  $d_j$  的输出节点. 这里,  $\| \cdot \|$  表示欧氏距离.

(4) 对节点  $j^*$  和邻域内的所有节点集合  $NE_j(t)$  按 UCL 算法修改突触向量

$$M_{ij}(t+1) = M_{ij}(t) + \eta(t)[X_j(t) - M_{ij}(t)] \quad (6)$$

其中,  $j \in NE_j(t)$ ,  $\eta(t)$  为随时间递减的函数,  $0 \leq \eta(t) < 1$ . 本系统取

$$\eta(t) = 0.2 \exp(-5t/t_0)$$

这里,  $t_0$  为总学习次数,  $t$  为实际学习次数.

(5) 转第 2 步, 重复之. 值得指出的是, 关于学习过程中的邻域半径选取方法, 本系统提供了二维和一维选取邻域的两种方法.

### 2.2.1 邻域二维收敛方法

该方法将输出节点在二维平面上排成方阵(图 7),因此该方法要求输出节点的数目为整数的平方。学习开始时邻域初始半径  $r_0$  一般取输出节点总数的平方根的一半。

假设对某一输入矢量,输出节点 A 的欧氏距离最小,这时所取邻域的样本为所有输出节点。其邻域收敛依次为  $N_A(t_0) \rightarrow N_A(t_1) \rightarrow N_A(t_2)$ ;如果 B 点为欧氏距离最小点,那么其邻域收敛顺序依次为  $N_B(t_0) \rightarrow N_B(t_1) \rightarrow N_B(t_2)$ 。

### 2.2.2 邻域的一维收敛方法

为消除输出节点数必须为整数平方的要求,我们提出了邻域半径的一维收敛方法。所谓一维收敛是指输出节点呈线性分布。与一维数组类似,对输出节点的数目没有要求。初始半径  $r_0$  一般选输出节点数的一半。假设中间点 A 为对应某输入矢量的欧氏距离的最小点,则其邻域收敛次序依次为  $N_A(t_0) \rightarrow N_A(t_1) \rightarrow \dots$ (图 8a)。如输出节点有 9 个,B 点的欧氏距离最小时,其邻域收敛过程为  $N_B(t_0) \rightarrow N_B(t_1) \rightarrow \dots$ (图 8b)。

在本系统中邻域半径  $r$  的收敛公式为  $r = r_0(1 - t/t_0)$

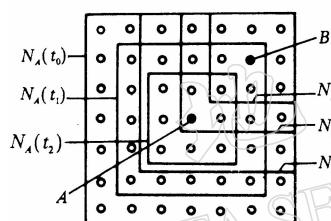


图 7 学习过程中的邻域半径二维收敛方法

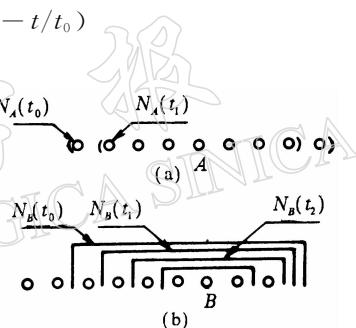


图 8 学习过程中的邻域半径一维收敛方法

## 3 结论与讨论

由于 FAM 模糊神经网络的使用,使得 NGESEP 成为国内外第一个具有学习功能的新一代地震预报专家系统。该系统通过学习得到了一批十分有价值的地震预报知识,一些知识是以往专家较难总结得到的。在实际进行学习时还需注意如下几点:

(1) 输入、输出语言值集合的划分应根据样本的群集情况,划分时要尽可能将群集的样本划在一个虚规则区域内。通常在地震样本不多的情况下,语言值集合的划分不必太细。

(2) 竞争突触向量数目  $q$  应大于虚规则数。通常  $q$  太大时,学习迭代时间长;  $q$  过小时,结论重复性差。实际学习时可取  $q=1\sim 2N$ ,这里  $N$  为学习样本数。

(3) 学习次数  $n$  对学习结果也有一定影响。 $n$  太大时,学习时间长; $n$  太小时,结论重复性差,通常应选取  $n\geq 1\,000$  次。

## 参 考 文 献

- 庄昆元,王炜,黄冰树,等,1991. 地震预报专家系统. 北京:地震出版社. 1~17  
 Kohonen T, 1984. *Self-Organization and Associative Memory*. Berlin: Springer-Verlag. 1~10  
 Kosko B, 1991. *Neural Network and Fuzzy System —— A Dynamical Systems Approach for Machine Intelligence*. Prentice Hall: Englewood Cliffs. 323~333