

一种基于最大-乘积型合成神经元的模糊联想记忆网络*

肖 平 余英林

(华南理工大学无线电与自动控制研究所·广州, 510641)

摘要: 本文提出了由最大-乘积型合成神经元的模糊神经网络实现双向联想记忆的一种学习方法及实现双向联想记忆的充要条件. 对于自联想记忆网络, 自联想权重由广义模糊解确定, 模式联想一次就可收敛, 该网络具有较强的容错性. 大量的计算机实验结果表明该学习算法是行之有效的.

关键词: 模糊神经网络; 模糊联想记忆; 广义模糊解

1 引 言

模糊系统是将一个模糊集映射到另一个模糊集合的系统, 其行为就象联想记忆一样, 把相近的输入映射成相近的输出, 由于信息的分布式存储, 联想记忆网络能够从不完整的或噪声的模式中联想出存储记忆中的某个完整的清晰的模式. 最大-乘积型模糊神经元突出的是主因素起决定作用, 即该神经元的总信息量等于各路输入中最大的那个, 这和逻辑“或”相对应. 特别在多层次综合决策问题中, 常常利用最大-乘积型算子作出判决决策, 在这两个算子中, 取大算子舍弃一些不重要的信息, 而乘积运算可体现规则之间的相关性, 同样可减少数据量, 而不致于对整个系统产生大的影响, 联想记忆网络的一个很重要的作用就是容错性能, 基于上述考虑, 我们较详细地研究了这种基于最大-乘积型合成神经元的网络结构, 包括学习算法, 收敛性等问题. 实验结果表明, 学习算法是行之有效的.

2 联想记忆网络结构与连接权学习算法

将普通的神经元的“乘积”与“和”算子分别用“ \vee ”和“ \cdot ”算子代替后得到的神经元称之为最大乘积型神经元, 由众多最大乘积型神经元连接起来或与其他神经元一起组成的网络叫做最大-乘积型神经网络. 一般地, 模糊神经网络的输入及权值是在 $[0, 1]$ 上取值, 输出函数 $f(x) = x$, 阈值 $\theta = 0$. 最大乘积型联想网络结构是一个较简单二层网络, 易于硬件实现.

最大-乘积型联想记忆网络的数学模型为

$$A_k \otimes W = B_k, \quad k = 1, 2, \dots, p, \quad (1)$$

即
$$\bigvee_{i=1}^n (a_{ki} \cdot w_{ij}) = b_{kj}, \quad j = 1, 2, \dots, m. \quad (2)$$

式中 \otimes 为 (\vee, \cdot) 型算子, $A_k = (a_{k1} \dots a_{kn}) \in [0, 1]^n$, $B_k = (b_{k1} \dots b_{km}) \in [0, 1]^m$. P 为记忆模式个数, 网络的连接权 $w_{ij} \in [0, 1]$.

记 $W_k = A_k \varepsilon B_k = (w_{ij}^k)_{n \times m}$ 为第 K 个训练模式对的编码矩阵, “ ε ” 为一个蕴涵算子. 定义

$$w_{ij}^k = \begin{cases} b_{kj}/a_{ki}, & \text{if } a_{ki} > b_{kj}, \\ 1, & \text{if } a_{ki} \leq b_{kj}. \end{cases} \quad (3)$$

一般来说, 连接权 W_k 只应由模式 A_k 到模式 B_k 的联想, 为了实现双向联想, 即除了满足 (1) 式外, 还要满足 $B_k \otimes M^T = A_k$. 可采用下面的反向编码方法. 设 $M = (m_{ij}^k)_{n \times m}$, 定义

* 国家“攀登计划”及国家自然科学基金资助项目.
本文于 1996 年 7 月 22 日收到, 1996 年 12 月 23 日收到修改稿.

$$m_{ij}^k = \begin{cases} a_{ki}/b_{kj}, & \text{if } a_{ki} > b_{kj} (b_{kj} \neq 0), \\ 0, & \text{if } a_{ki} < b_{kj}, \\ 1, & \text{else.} \end{cases} \quad (4)$$

于是对于每个模式对 (A_k, B_k) , 采用式(3)的编码方法, 得到 p 个模糊联想矩阵 W_k . 用最小叠加运算组合各个 W_k , 即

$$W = \bigcap_{k=1}^p W_k = \bigcap_{k=1}^p A_k^T \varepsilon \cdot B_k. \quad (5)$$

其中“ \cap ”为模糊交运算. (5)式用逐点表示, 即为

$$w_{ij} = \bigwedge_{k=1}^p w_{ij}^k = \bigwedge_{k=1}^p \{b_{kj}/a_{ki} | b_{kj} \leq a_{ki}\}. \quad (6)$$

同理, 利用最小叠加运算把 P 个反向联想矩阵组合起来, 即

$$m_{ij} = \bigwedge_{k=1}^p m_{ij}^k = \bigwedge_{k=1}^p \{a_{ki}/b_{kj} | b_{kj} \leq a_{ki}\}. \quad (7)$$

定理 2.1 若 $S(W) = \{W | A_k \otimes W = B_k, k \in \Omega\} \neq \emptyset$, 则由(5)式所确定的连接权矩阵 $W = (w_{ij})_{n \times m}$ 是 $S(W)$ 中的最大模糊连接权矩阵.

定理 2.2 给定存储模式对 $(A_k, B_k) (k = 1, 2, \dots, p)$, 则存在连接权矩阵 W , 使对于一切 $k \in \Omega$, $A_k \otimes W = B_k$ 的充要条件是

$$\bigcup_{i=1}^n S_{ij} = \bigcup_{i=1}^n \{k \in \Omega | b_{kj}/a_{ki} \leq w_{ij}, b_{kj} \leq a_{ki}\} = \Omega, \quad j = 1, 2, \dots, m. \quad (8)$$

定理 2.3 给定存储模式对 (A_k, B_k) . $L(A_k) \geq L(B_k) (k = 1, 2, \dots, p)$, 则存在连接权矩阵 M , 使对于一切 $k \in \Omega$, $B_k \otimes M^T = A_k$ 的充要条件是

$$\bigcup_{j=1}^m S_{ij} = \bigcup_{j=1}^m \{k \in \Omega | a_{ki}/b_{kj} \leq m_{ji}, b_{kj} \leq a_{ki}\} = \Omega, \quad i = 1, 2, \dots, n. \quad (9)$$

其中 $L(A_k) = \min \{a_{ki} | i = 1, 2, \dots, n\}$, $L(B_k) = \min \{b_{kj} | j = 1, 2, \dots, m\}$.

推论 2.1 对于每一个模式对 (A_k, B_k) , 则 $A_k \otimes W = B_k$ 的必要条件是 $H(A_k) \geq H(B_k)$. 其中 $H(A_k) = \max \{a_{ki} | i = 1, 2, \dots, n\}$, $H(B_k) = \max \{b_{kj} | j = 1, 2, \dots, m\}$.

推论 2.2 设 A'_k 是模式 A_k 丢失部分信息所产生的模式, 且 $H(A'_k) = H(A_k)$. 若 $A_k \otimes W = B_k$ 则 $A'_k \otimes W = B_k$.

需要指出的是上述反向编码矩阵 M 的元素值并不在 $[0, 1]$ 中, 但它确实是模糊方程 $B_k \otimes M^T = A_k$ 的一个解, 我们称其为模糊方程 $B_k \otimes M^T = A_k$ 的广义模糊解.

很显然, 由定理 2.2 和定理 2.3 不难得到模式对 $(A_k, B_k) (k \in \Omega)$ 可以双向联想的充要条件是(8), (9)同时成立. 对于自联想模式对 $(A_k = B_k, \text{即 } a_{ki} = b_{ki})$ 易知 $w_{ii} = m_{ii} = 1$. 于是对于任一个 $k \in \Omega$, 都有 $b_{ki}/a_{ki} = a_{ki}/b_{ki} = 1 = w_{ii} = m_{ii}$, 故 $\bigcup_{j=1}^m S_{ij} = \Omega$. 条件(8), (9)自然满足. 因此任一自联想模式均为网络的稳定态, 换句话说, 最大-乘积型联想网络可以记住任意多个自联想模式对. 本文我们主要研究自联想记忆网络, 对异联想记忆网络, 我们将进一步研究.

3 自联想网络连接权矩阵的学习算法

对于自联想模式对, 虽然 $A_k \otimes W = A_k$ 或 $A_k \otimes M^T = A_k (k \in \Omega)$, 但以 W 或 M^T 作为自联想网络的权矩阵并不是最好的, 而以

$$R = \bigwedge_{k=1}^p R_k = \bigwedge_{k=1}^p (W_k \otimes M_k^T) \quad (R_k = W_k \otimes M_k^T) \quad (10)$$

作为网络的连接权更恰当, 这只要比较下面定理 3.1 与推论 2.2 即可看出.

定理 3.1 设 A'_k 是模式 A_k 丢失部分信息所产生的模式, 且 $H(A_k \cap A'_k) \neq 0$ 若 $A_k \otimes (W$

$\otimes M^T) = A_k$, 则 $A'_k \otimes (W \otimes M^T) = A_k$. 这里 W, M^T 分别由(3)与(4)确定.

通过简单的计算可知 $W \otimes M^T = R = (r_{ij})_{n \times n}$ 中的元素 $r_{ii} = 1$ (R 满足自反性), $r_{ij} = 1/r_{ji}$ ($r_{ji}, r_{ij} \neq 0$), 因此有些 $r_{ij} \in [0, 1]$, 但它确是模糊方程 $A \otimes X = A$ 的一个广义模糊解, 并且 R 还是传递的, 即 $R \otimes R = R$. 由此不难得到

定理 3.2 设 $A_k \otimes W = A_k, A_k \otimes M^T = A_k. (k = 1, 2, \dots, p)$ 则由(10)式确定的 R 满足 $A_1 \otimes R = A_1 \otimes \bigwedge_{k=1}^p (W_k \otimes M_k^T) = A_1 (l = 1, 2, \dots, p)$ 且 $R \otimes R = R$.

定理 3.2 说明本文给出的自联想记忆网络对于任一初始状态模式 A 只需要循环一次即可收敛到网络的另一个稳定态. 这是因为任一模式 A 的联想过程可表示为 $A \rightarrow W \rightarrow A' \rightarrow M^T \rightarrow A'' \rightarrow \dots \rightarrow A_f \rightarrow \dots$ 由定理 3.2 易知, $A \otimes R = A \otimes R^2 = \dots = A \otimes R^n = A_f$, 因此, $A \otimes R \otimes W = A_f \otimes W$. 这也说明 R 作为最大-乘积型自联想网络的连接权是全局稳定的.

4 模拟实验

实验 1 实验的目的是考察最大-乘积型联想网络对于自联想模式部分信息丢失的联想能力. 图 1 是一幅 64×64 , 灰度级为 256 的 Lenna 图象的头部, 我们分别以输入层节点数为 128 和输入层节点数为 256 的两个联想网络进行实验. 因此, 在第一个网络中, 图 1 被分成 32 个样本模式(每个模式为 128 个象素值), 在第二个网络中, 图 1 被分成 16 个样本模式(每个模式为 256 个象素值), 广义模糊解 R 作为网络连接权. 图 2 是图 1 中丢失了小于 110 的象素值, 并随机丢失 160 个其它象素值(图 2 的黑色点就是丢失的数据点), 共约丢失全部数据的 30% ($1209/64 \times 64$). 图 3, 图 4 分别是第一个网络和第二个网络对图 2 联想恢复的结果, 它们各自的信噪比列于表 1. 实验中可看出, 随着记忆模式数的增加, 网络的容错性将变差.



图 1 Lenna 头像 图 2 噪声图像 图 3 网络 1 联想结果 图 4 网络工联想结果

由于最大乘积型模糊算子突出的是主因素作用, 因此部分信息丢失或模式分量值较小, 网络可以较好的联想, 而对于另一种噪声, 即模式分量值变大, 尤其是脉冲性噪声, 网络将难以进行完整联想. 但只要采用下面的预处理

表 1 峰值信噪比

名称	图 2	图 3	图 4
PSNR(峰值信噪比)	16.40	32.18	40.39

理方式, 即在最大-乘积型网络前加一个预处理层如 Maxnet 网络或 Hamming 网络^[4]. 下面两个实验均是以这种方法处理的. 具体方法概括为: 设 $C = (c_1, c_2, \dots, c_n) \in [0, 1]^n$ 是一个含噪声的模式. $A_i = (a_{i1}, a_{i2}, \dots, a_{in}) \in [0, 1]^n$ 是 P 个记忆模式 ($i = 1, 2, \dots, p$).

- 1) 计算 C 与所有 A_i 的欧氏距离或海明距离 $H(A_i, C)$, 记 $H(A_k, C) = \min H(A_i, C)$.
- 2) 比较 C 与 A_k 的对应分量的大小, 得 $B = C \wedge A_k = (c_1 \wedge a_{k1}, c_2 \wedge a_{k2}, \dots, c_n \wedge a_{kn})$.
- 3) 把模式 B 输入给网络进行联想.

上述处理方法的理由及作用是: 一方面因为任一输入模式总是收敛于与其最相近的特征

模式,而收敛模式的吸引域大小可用欧氏距离或 Hamming 距离来刻画^[2],另一方面, $C \wedge A$ 的作用可以纠正模式分量值由小变大的噪声点.

实验 2 实验的目的是综合考察最大乘积型联想网络随着自联想模式数目的增加,网络的容错能力.我们随机产生分量个数为 64,分量值大小在 0~255 之间的 P 个模式,对于每一个模式随机产生 50 个加一定幅度的高斯噪声模式作为测试样本,采用欧氏贴近度^[5]比较结果状态模式与期望的联想模式,若贴近度在 0.95 以上,称该噪声模式联想成功,如果全部测试模式的 98% 都联想成功,我们就说网络能够容错这种幅度的高斯噪声,并能记住 P 个模式.图 5 中横坐标表示记忆数目,纵坐标是

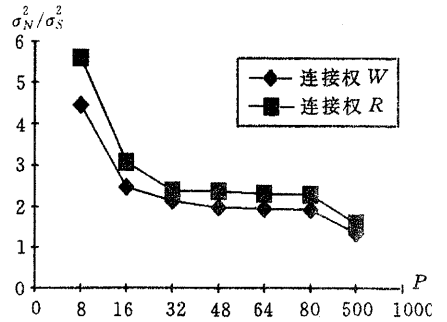


图 5 网络容错性能

高斯噪声方差 σ_N^2 除以输入与输出模式的差值 $\sigma_S^2 = \frac{1}{64} \sum_{k=1}^P (\bar{b}_k - b_k)^2$, (\bar{b}_k 是实际联想状态模式, b_k 是期望联想模式. 实验中 σ_S^2 取所有联想成功的输入与输出模式的差值的平均值), 由图 1 可以看出, 由广义模糊解 R 确定的连接权好于最大模糊权 W 确定的学习算法. 当记忆模式数为 500 时, 网络仍有较强的容错能力. 当随着网络记忆的模式数的增加, 网络的容错性能变差, 但能够记住所有的模式.

实验 3 为了比较 Hopfield 网络与本文网络对于二值状态模式(以 0.1 和 0.9 表示)吸引域的大小, 我们重复做了文[2]的实验. 训练样本是 10 个字符(如图 6 所示), 采用噪声输入测试网络吸引域大小, 噪声输入是由训练样本 A_i 随机反转(0.1 变 0.9 或 0.9 变 0.1)某些神经

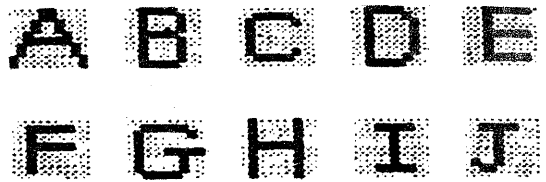


图 6 待记忆的字符样本

元状态产生的. 反转率从 0.0 增大到 0.4, 对于每一个反转率 R , 测试 5000 个噪声模式. 若噪声模式作为网络初始状态, 系统能收敛于相应的样本 A_i , 则称为一次成功, 否则是失败. 表 2 是联想成功率. Hopfield 网络的学习权重采用目前三种较有效果的学习算法^[3,7,8], 可以看出本文的模糊网络的学习算法形成的吸引域较大.

表 2 联想成功率

反转率	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
本文网络	100%	100%	100%	100%	100%	98%	82%	77%	67%
约束最小化	100%	100%	100%	100%	100%	95%	82%	66%	32%
Cardner	100%	100%	100%	98%	95%	89%	71%	45%	19%
最小重叠	100%	100%	100%	99%	98%	91%	74%	51%	18%

5 结束语

本文提出了最大-乘积型模糊联想记忆网络及其学习算法, 从理论上详细分析了自联想及异联想实现的条件, 并对自联想记忆网络的应用作了一些探索. 实验证实了结论的正确性及学习算法的有效性.

参 考 文 献

- 1 范俊波等. 模糊联想记忆的一种有效学习算法. 电子学报, 1996, 24(1): 112-115
- 2 汪涛等. 联想记忆网络的约束化学习. 计算机学报, 1995, 16(12): 11-13
- 3 张承福等. 联想记忆神经网络的若干问题. 自动化学报, 1994, 20(5): 19-21
- 4 周继成. 人工神经网络. 北京: 科学普及出版社, 1993
- 5 李晓忠等. 模糊神经网络. 贵阳: 贵州科技出版社, 1994
- 6 罗承忠等. 模糊集引论(上册). 北京: 北京师范大学出版社, 1989
- 7 Carden. The Space of interaction in neural network models. J. Phys. A: Math. Gen., 1988, 21: 257-270
- 8 Krauth, W. Mezard, M.. Learning algorithm with optimal stability in neural networks. J. Phys. A: Math. Gen., 1987, 20: 745-752

A Kind of Fuzzy Associative Memories Networks Based on Max-Product Composition Units

XIAO Ping and YU Yinglin

(The research institute of Radio and Automation, South China University of Technology • Guangzhou, 510641, PRC)

Abstract: A kind of fuzzy associative memories and its learning algorithm with max-product composition units is presented in this paper. The necessary and sufficient condition to realise bidirectional associative memories is given. The patterns for auto-memories can be convergenced through the connective weight matrix for iteration just once. The connection weight matrix for auto-memories is determined by general fuzzy solution. The computer simulation results have shown the efficiency of the fuzzy network and its learning algorithm.

Key words: neural network; fuzzy associative memory; connection weight matrix

本文作者简介

肖平 1964年生. 1983年毕业于赣南师院数学系. 1988年获国防科技大学应用数学专业硕士学位. 现为华南理工大学博士研究生. 主要研究方向是模糊数学, 神经网络, 图象处理, 已发表论文十余篇.

余英林 1953年毕业于华南工学院. 1961年12月中科院电子所副博士研究生毕业. 现为华南理工大学无线电所博士生导师. 从事图象处理, 模式识别, 神经网络方面的研究.