文章编号:1001 - 9081(2007)02 - 0346 - 03

训练模式的摄动对最大—乘积型模糊联想记忆网络的影响

曾水玲[1,2],徐蔚鸿[1,2],杨静宇[3]
(1. 吉首大学 数学与计算机科学学院,湖南 吉首 416000;
2. 长沙理工大学 计算机与通信工程学院,湖南 长沙 410077;
3. 南京理工大学 计算机科学与技术学院,江苏 南京 416000)
(zengflsl@163.com)

摘 要:首先建立了前馈型模糊联想记忆网络对训练模式摄动的鲁棒性概念,分析了最大—乘积型模糊联想记忆网络(Max-Product FAM),发现当采用模糊赫布学习算法时它的鲁棒性好,但采用另一学习算法时鲁棒性较差。最后用实验验证了理论结果。

关键词:模糊联想记忆网络;学习算法;训练模式对;摄动;鲁棒性

中图分类号: TP18 **文献标识码**:A

Influences of perturbations of training pattern pairs on max-product fuzzy associative memory

ZENG Shui-ling[1,2], XU Wei-hong [1,2], YANG Jing-yu[3]
(1. *College of Mathematics and Computer Science, Jishou University, Jishou Hunan* 416000, *China*;
2. *College of Computer and Communications Engineering, Changsha University of Science and Technology, Changsha Hunan*410077, *China*;
3. *College of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing Jiangsu* 210094, *China*)

Abstract: A new concept was established in this paper that the robustness of a feed-forward fuzzy associative memory to perturbations of training pattern pair. Then a Max-Product-based Fuzzy Associative Memory (Max-Product FAM) was analyzed. Investigation reveals that such robustness of the memory is good when the fuzzy Hebbian learning algorithm is used, however is poor when another learning algorithm is employed. Finally, an experiment is given to testify the theoretical conclusion and illustrate practical application of Max-product FAM.

Key words: fuzzy associative memory network; learning algorithm; pattern pair; perturbation; robustness

0 引言

模糊神经网络是人工神经网络和模糊逻辑的互补式结合[1~3],能对人脑的生物结构、功能以及信息处理的过程和特点进行综合模拟,集学习、联想、识别、自适应及模糊信息处理于一体。它的研究范畴大体上分为:新模型设计,性能分析(存储容量、逼近能力、容错性、鲁棒性、反馈网络的收敛性和稳定性等),学习算法设计,硬件实现,以及在科学和工程领域的实际应用。

正如我们所知,构建一个实际的模糊系统通常需要准备描述系统性能的先验(前期)知识。如,模糊推理系统的已知规则,模糊神经网络系统的训练模式。而人们得到的先验知识一般含有少量误差,是现实中真实信息摄动的结果。为此,文献[4~8]从不同角度研究了已知规则摄动对模糊推理系统的影响。文献[9]结合学习算法研究训练模式摄动对模糊神经网络系统性能的影响及其控制问题,建立了这类问题的研究框架,并以模糊双向联想记忆网络(Max-Min FBAM)为例进行了具体分析。但该文对前馈型模糊联想记忆网络尚未给出分析实例,特别是文中的定义 3 在刻画网络的输出摄动时适宜采用更一般化的描述。本文将首先就前馈型模糊联想记忆网络修改文献[9]中的定义 3,接着具体研究最大—乘积型联想记忆网络(Max-Product FAM)对训练模式摄动的鲁棒性问题(此时网络分别采用糊赫布学习以及更好的最大—最小编码学习算法[10]进行训练),以期为类型和学习算法都具有多样性的前馈型模糊联想记忆网络的此类实用问题的研究进行尝试。

1 相关定义和引理

本文假定 $I = \{1,2,\cdots,n\}$,$J = \{1,2,\cdots,m\}$ 为非空有限指标集。设 $A = (a_1,a_2,\cdots,a_n) \in [0,1]^n$,$\Delta A = (\Delta a_1,\Delta a_2,\cdots,\Delta a_n) \in [-1,1]^n$,定义 $A + \Delta A = (a_1 + \Delta a_1, a_2 + \Delta a_2,\cdots, a_n + \Delta a_n) \in [0,1]^n$,其中,若 $a_i + \Delta a_i > 1$,则 $a_i + \Delta a_i$ 调整为 1,若 $a_i + \Delta a_i < 0$,则 $a_i + \Delta a_i$ 调整为 0。其余类似情形可类似定义,此处略。

令 $A = (a_1,a_2,\cdots,a_n) \in [0,1]^n$,$B = (b_1,b_2,\cdots,b_m) \in [0,1]^m$,称模糊模式对 (A,B) 中的 A 为模式对的前件,B 为模式对的后件。

收稿日期:2006 - 08 - 29;**修订日期**:2006 - 11 - 20
基金项目:国家自然科学基金资助项目(60474070);湖南省自然科学基金资助项目(05JJ40004)
作者简介:曾水玲(1975 -),女,湖南湘潭人,讲师,硕士研究生,主要研究方向:神经网络、模式识别; 徐蔚鸿(1963 -),男,湖南湘潭人,教授,博士,主要研究方向:智能系统、模式识别、软件工程等; 杨静宇(1941 -),男,江苏南京人,教授,博士生导师,主要研究方向:模式识别、智能系统、图像处理、机器人.

定义 1[9]　设 $A^*, A \in [0,1]^n$,称 $W(A^*,A) = \bigvee_{i\in I} | a_i^* - a_i |$ 为 A^* 与 A 的最大摄动误差。

定义 2[9]　当模式对 (A,B) 被变为 $(A+\Delta A, B)$,且 $W(A+\Delta A, B) \leqslant \varepsilon$ 时,称该模式对发生了前件最大 ε 摄动;当模式对 (A,B) 被变为 $(A, B+\Delta B)$,且 $W(A, B+\Delta B) \leqslant \varepsilon$ 时,称模式对发生了后件最大 ε 摄动;当规则 (A,B) 被变为 $(A+\Delta A, B+\Delta B)$ 且 $W(A+\Delta A, A) \vee W(B+\Delta B, B) \leqslant \varepsilon$ 时,称该模式对发生了最大 ε 摄动。

设现实中客观的模式对为 (C,D),对其用某种方法测量后获得的结果为 (A,B),则显然 (A,B) 和 (C,D) 彼此可看成是对方小幅摄动的结果。

一个前馈型模糊联想记忆网络基于训练模式对集 *training_set*,采用学习算法 f 进行训练后,当输入 $A' \in [0,1]^n$ 时,网络产生的输出记为 $f(A', training\_set) \in [0,1]^m$。

定义 3　设一个前馈型模糊联想记忆网络采用学习算法 f,若存在常数 $h > 0$,使得对任意摄动幅度 $\gamma_k \in (0,1)$ 和任意给定的训练模式对集合 $set = \{(A_k, B_k) \mid k = 1,2,\cdots,p\}$,当各模式对 (A_k, B_i) 发生最大 γ_k 摄动而把 *set* 变成新训练模式对集合 *new_set*,若对一切输入 $A^* \in [0,1]^n$,有 $W(f(A^*, set), f(A^*, new\_set)) \leqslant h \bigvee_{k=1}^{p} \gamma_k$ 总成立,则称采用学习算法 f 的前馈型模糊联想记忆网络对训练模式集摄动幅度在系数为 h 的条件下全局不放大。

值得强调的是定义 3 中:1)*set* 中的模式的构成和个数 p 都是任意的;2)输入 A^* 是任意的;3)$\bigvee_{k=1}^{p} \gamma_k$ 也可看成是整个模式集 *set* 的最大摄动幅度;4)此定义中,把文献[9]的定义 3 中 $W(f(A^*, set), f(A^*, new\_set)) \leqslant \bigvee_{k=1}^{p} \gamma_k$ 改成了 $W(f(A^*, set), f(A^*, new\_set)) \leqslant h \bigvee_{k=1}^{p} \gamma_k$。

引理 1[9]　$\left| \bigvee_{i\in I} a_i' - \bigvee_{i\in I} a_i \right| \leqslant \bigvee_{i\in I} | a_i' - a_i |$。

2　对训练模式和摄动的鲁棒性分析

文献[2]中提出的最大 — 乘积型 FAM 网络的数学模型为:$X \otimes W = Y$,可逐点表示为:$\bigvee_{i=I}^{n} (x_i \cdot w_{ij}) = y_j, \forall j \in J$。

其中 $\otimes$ 为最大—乘积型算子,$X = (x_1, x_2, \cdots, x_n) \in [0,1]^n$,$Y = (y_1, y_2, \cdots, y_n) \in [0,1]^m$,$\boldsymbol{W} = (w_{ij})_{n\times m}$ 为神经网络的连接权矩阵,$w_{ij} \in [0,1]$ 为第 i 个输入神经元到第 j 个输出神经元之间的连接权。

2.1　采用模糊赫布规则训练网络

模糊赫布规则学习算法[2]首先采用乘积算子将第 k 个模式对 (A_k, B_k) 编码到矩阵 $\boldsymbol{W}_k$,即 $\boldsymbol{W}_k = (a_i^k \cdot b_j^k)_{n\times m}$,然后用最大叠加运算将这 p 个 $\boldsymbol{W}_k$ 组合到最终的连接权矩阵 $\boldsymbol{W} = \bigcup_{k=1}^{p} \boldsymbol{W}_k$,即 $\boldsymbol{W} = (w_{ij})_{n\times m} = (\bigvee_{k=I}^{p} (a_i^k \cdot b_j^k))_{n\times m}$。

定理 1　采用模糊赫布规则学习算法时,Max-Product FAM 对训练模式集摄动幅度在系数为 2 的条件下全局不放大。

证明:任给训练模式集 $set = \{(A_k, B_k) \mid k = 1,2,\cdots,p\}$,采用模糊赫布规则学习算法对模糊联想记忆网络训练完成后,得到权值矩阵 $\boldsymbol{W}$,让网络开始工作。

任意给定输入 $A^* \in [0,1]^n$,由网络得到的输出为 $B^* = A \otimes \boldsymbol{W}$,用逐点方式可表为 $b_j^* = \bigvee_{i=I}^{n} (a_i^* \cdot w_{ij}) = \bigvee_{i=I}^{n} [a_i^* \cdot (\bigvee_{k=I}^{p} (a_i^k \cdot b_j^k))], j \in J$。

假定训练模式集 *set* 摄动后变为 $set\_set = \{(C_k, D_k) \mid k = 1,2,\cdots,p\}$,其中 $W(A_k, C_k) \leqslant \varepsilon_k$,$W(B_k, D_k) \leqslant \beta_k$,$k = 1,2,\cdots,p$,也采用模糊赫布规则学习算法,训练网络得到权值矩阵 $\boldsymbol{W}1$,由于训练模式集的变化,此时网络的权值会有某些变化,即一般地 $\boldsymbol{W} \neq \boldsymbol{W}1$,现让网络开始工作。对同一输入 A^*,由网络得到的输出为 $B1^* = A^* \otimes \boldsymbol{W}1$,即 $b1_j^* = \bigvee_{i=1}^{n} (a_i^* \cdot w1_{ij}) = \bigvee_{i=I}^{n} [a_i^* \cdot (\bigvee_{k=1}^{p} (c_i^k \cdot d_j^k))], j \in J$。

$$
\begin{aligned}
| b_j^* - b1_j^* | &= \left| \bigvee_{i=1}^{n} [a_i^* \cdot (\bigvee_{k=1}^{p} (a_i^k \cdot b_j^k))] - \bigvee_{i=1}^{n} [a_i^* \cdot (\bigvee_{k=1}^{p} (c_i^k \cdot d_j^k))] \right| \\
&\leqslant \bigvee_{i=1}^{n} \left| [a_i^* \cdot (\bigvee_{k=1}^{p} (a_i^k \cdot b_j^k))] - [a_i^* \cdot (\bigvee_{k=1}^{p} (c_i^k \cdot d_j^k))] \right| \quad (\text{由引理 1}) \\
&\leqslant \bigvee_{i=1}^{n} \left| (\bigvee_{k=1}^{p} (a_i^k \cdot b_j^k)) - (\bigvee_{k=1}^{p} (c_i^k \cdot d_j^k)) \right| \quad (\text{由 } a_i^* \leqslant 1) \\
&\leqslant \bigvee_{i=1}^{n} \bigvee_{k=1}^{p} | (a_i^k \cdot b_j^k) - (c_i^k \cdot d_j^k) | \quad (\text{由引理 1}) \\
&\leqslant \bigvee_{i=1}^{n} \bigvee_{k=1}^{p} (| a_i^k b_j^k - a_i^k d_j^k + a_i^k d_j^k - c_i^k d_j^k |) \\
&\leqslant \bigvee_{i=1}^{n} \bigvee_{k=1}^{p} [| a_i^k || b_j^k - d_j^k | + | d_j^k || a_i^k - c_i^k |] \\
&\leqslant \bigvee_{i=1}^{n} \bigvee_{k=1}^{p} [| b_j^k - d_j^k | + | a_i^k - c_i^k |] \quad (| a_i^k | \leqslant 1, | d_j^k | \leqslant 1) \\
&\leqslant \bigvee_{i=1}^{n} \bigvee_{k=1}^{p} (\varepsilon_k + \beta_k) \\
&\leqslant \bigvee_{i=1}^{n} \bigvee_{k=1}^{p} 2(\varepsilon_k \vee \beta_k) = \bigvee_{i=1}^{n} (\bigvee_{k=1}^{p} 2\gamma_k) = 2 \bigvee_{k=1}^{p} \gamma_k
\end{aligned}
$$

故 $W(B^*, B1^*) = \bigvee_{j=1}^{m} | b_j^* - b1_j^* | \leqslant 2 \bigvee_{k=1}^{p} \gamma_k$,依定义 3 则有 Max − Product FAM 对训练模式集最大摄动幅度在系数为 $h = 2$ 的条件下全局不放大。证毕。

2.2　采用文献[10]的学习算法进行训练网络

文献[10]为 Max-Product FAM 提出的学习算法是:由第 k 个模式对 (A_k, B_k) 编码到中间矩阵 $\boldsymbol{W}_k = (w_{ij}^k)_{n\times m} = (a_i^k \theta b_j^k)_{n\times m}$ 中,然后采用最小叠加运算将所有 p 个 $\boldsymbol{W}_k$ 组合到最终的连接权矩阵 $\boldsymbol{W} = \bigcap_{k=1}^{p} \boldsymbol{W}_k$ 中,即 $\boldsymbol{W} = (w_{ij})_{n\times m} = (\bigwedge_{k=1}^{p} w_{ij}^k)_{n\times m} = (\bigwedge_{k=1}^{p} (a_i^k \theta b_j^k))_{n\times m}$。

$$
\text{其中 } w_{ij}^k = a_i^k \theta b_j^k = \begin{cases} 1, & a_i^k = b_j^k \\ \dfrac{\min(a_i^k, b_j^k)}{\max(a_i^k, b_j^k)}, & a_i^k \neq b_j^k \end{cases}
$$

定理 2　采用以上学习算法时,在任意正常数 h 的条件下,Max-Product FAM 不具有对训练模式集摄动幅度全局不放大的性质。

证明:依定义 3,如果能证明:对 $\forall h > 0$,总存在某个 $\gamma \in (0,1)$ 以及某一训练模式集 set_0 使得其中各模式对发生最大 γ 摄动而变成 new_set_0 时,总 $\exists A' \in [0,1]^n$ 使得 $W(f(A', set_0), f(A', new\_set_0)) > h\gamma$ 成立,则就证明了定理 2 成立。

事实上,对任意给定的 $h > 0$,若 h 满足 $2 > h > 0$,取 γ 满足 $0 < \gamma < \min\{1/h, 1/(2-h), 1\}$,若 $h \geqslant 2$,取 γ 满足 $0 < \gamma < 1/h$。又取 $t \in (0, (1-\gamma h)/(1+\gamma h)) \subset (0,1)$,这样总有 $\gamma h \in (0,1)$,$t^2\gamma, t\gamma, (t+1)\gamma \in (0,1]$ 都成立。并构造 $set_0 = \{(A_1, B_1)\} = \{((t\gamma, t\gamma, \cdots, t\gamma)_n, (t^2\gamma, t^2\gamma, \cdots, t^2\gamma)_m)\}$,其发生最大 γ 摄动而变成:

$new\_set_0 = \{(C_1, D_1)\} = \{(((t+1)\gamma, \cdots, (t+1)\gamma),$

$((t^2+1)\gamma,\cdots,(t^2+1)\gamma))\}$,接着取特定输入$A'=(1,1,\cdots,1)_n$。

$$\forall j \in J,$$

$$\begin{aligned}
|b_j' - b1_j'| &= \left|\bigvee_{i=1}^{n} a_i' \cdot w_{ij} - \bigvee_{i=1}^{n} a_i' \cdot w1_{ij}\right| \\
&= \left|\bigvee_{i=1}^{n} [a_i' \cdot \bigwedge_{k=1}^{p} (a_i^k \theta b_j^k)] - \right. \\
&\quad \left. \bigvee_{i=1}^{n} [a_i' \cdot \bigwedge_{k=1}^{p} (c_i^k \theta d_j^k)]\right| \\
&= \left|\bigvee_{i=1}^{n} [1 \cdot \bigwedge_{k=1}^{1} (a_i^k \theta b_j^k)] - \bigvee_{i=1}^{n} [1 \cdot \bigwedge_{k=1}^{p} (c_i^k \theta d_j^k)]\right| \\
&= \left|\bigvee_{i=1}^{n} (a_i^1 \theta b_j^1) - \bigvee_{i=1}^{n} (c_i^1 \theta d_j^1)\right| \\
&= \left|\bigvee_{i=1}^{n} \frac{t^2\gamma}{t\gamma} - \bigvee_{i=1}^{n} \frac{(t^2+1)\gamma}{(t+1)\gamma}\right| \quad (\text{依 } \theta \text{ 的定义}) \\
&= \left|t - \frac{t^2+1}{t+1}\right| \\
&= \frac{2}{t+1} - 1 > rh \\
&\qquad (\text{因 } t < (1-\gamma h)/(1+\gamma h))
\end{aligned}$$

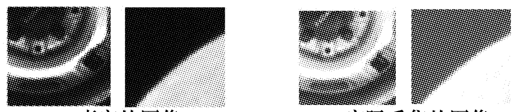
又 $W(f(A',set_0),f(A',new\_set_0)) = \bigvee_{j=1}^{m} |b_j' - b1_j'| \geqslant \bigvee_{j=1}^{m} \gamma h = \gamma h$。由 $h > 0$ 的任意性和定义 3,知定理 2 的结论成立。

3 模拟实验

实验目的:1)对 $h=2$ 时定理 2 的结论进行验证;2)给出 Max-Product FAM 在图像联想方面的实际应用。

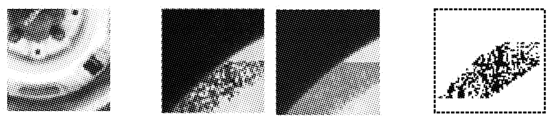
实验步骤:

1)图 1 中两幅灰度级为 256、大小为 64×64 的图像,不妨看作真实的图像。把图 1(a)的两幅图像的灰度值作不等量的小幅变大(在视觉上变白)得到图 1(b),不仿把图 1(b)中的图像看作是图 1(a)中的图像实际采集到的结果。



(a)真实的图像 (b)实际采集的图像

图 1 实验图像



(a)输入图像 (b)由FAM1和FAM2输出结果构成的图像 (c)两输出结果的比较

图 2 实验结果

2)图 1 中(a)图像的上半部分(含 32×64 个像素点)分别作为一个训练模式对的前件和后件,由左、右图像的下半部分构成另一训练模式对,并构建一个分别含有 32×64 个输入、输出节点的 Max-Product FAM1。类似地由图 1(b),能得到另外的两个训练模式对和 Max-Product FAM2,并采用文献[10]的学习算法训练两网络。

3)将图1(a)的左图像灰度值作另一小幅变大得到图2(a)。把图 2(a)的上、下两部分作为输入数据,依次送入 Max-Product FAM1,所得到的两输出构成图 2(b)中的左图像;再依次送入 Max-Product FAM2,所得到对应图像为图 2(b)中的右图像。

4)比较图 2(b)中的两图像的差异,用可视化形式表示成图 2(c),其中白色点(黑色点)表示两图像对应位置像素值的差异不大于(大于)两训练模式对的最大摄动幅度的 h 倍。

从图 2(c)可以看到,当使用文献[10]提出的学习算法时,模糊联想记忆网络不具有对任意训练模式集最大摄动幅度全局不放大的性质。在实验中我们采用的是同一网络模型、同一学习算法和相同的输入,但是两个网络产生的输出差异很大,这说明是由训练模式对的摄动导致的。

4 结语

本文对 Max-Product FAM 的训练模式获取过程能提供警示,对它的多种学习算法的选择以及性能分析有一定指导意义。进一步研究的内容包括:1)采用学习算法时,Max-Product FAM 不具有对训练模式集最大摄动幅度全局不放大的性质,若能对训练模式集的摄动给出某种可行的制约,使得在此条件下该网络具有对训练模式集最大摄动幅度全局不放大的性质。对其他的 t-模 T,Max-T FAM 受训练模式集摄动的影响可作类似的分析。2)对于一个采用学习算法 f 的前馈型模糊联想记忆网络,当其改造成相应的双向模糊联想记忆网络时,而学习算法不变,它们对于训练模式摄动的鲁棒性有什么差异。

参考文献:

[1] KOSKO B. Fuzzy associative memories[A]. In: Kandel A(ed): Fuzzy Expert Systems Reading[C]. MA, USA: Addison Wesley, 1987.

[2] KOSKO B. Neural Network and Fuzzy system[M]. New jersey, USA: Prentic-Hall, Inc., 1992, 24(1): 112-114.

[3] 王士同. 神经模糊系统及其应用[M]. 北京: 北京航空航天大学, 1998.

[4] YING MS. Perturbation of fuzzy reasoning[J]. IEEE Transactions on Fuzzy systems, 1999, 7(5): 625-629.

[5] CAI K-Y. Rubstness of Fuzzy Reasoning and -δEquations of Fuzzy Sets[A]. IEEE Trans. on Fuzzy Systems[C]. 2001. 738-750.

[6] ZHANG L, CAI K-Y. Optimal Fuzzy Reasoning and Its Robustness Analysis[A]. Int. J. of Intelligent Systems[C]. 2004. 1033-1049.

[7] LI YM, LI DC, PEDYCZ W, *et al*. An Approach to Measure the Robustness of Fuzzy Reasoning[A]. Int. J. of Intelligent Systems[C]. 2005. 393-413.

[8] 徐蔚鸿, 陈国平, 杨静宇, 等. 规则摄动时模糊蕴涵算子对模糊推理的鲁棒性的影响[J]. 计算机学报, 2005, 28(10): 1700-1707.

[9] 徐蔚鸿, 宋鸾姣, 李爱华, 等. 训练模式对的摄动对模糊双向联想记忆网络的影响及其控制. 计算机学报, 2006, 25(2) 153-157.

[10] 肖平, 杨丰, 余英林. 最大—乘积型模糊联想记忆网络的最大最小编码学习算法[J]. 通信学报, 1999, 20(1): 17-22.

(上接第 339 页)

[4] 刘琛. 连续小波变换在生物医学信号处理中的应用[J]. 重庆大学学报, 2003, 26(8): 23-26.

[5] HAYKIN S. Neural networks: A Comprehensive Foun-dation (Second Edition) [M]. Beijing: Tinghua University Press, 2001, ISBN7-302-04936-X/TP 2778. 318-350.

[6] VAPNIK VN. The Nature of Statistical Learning Theory [M]. New York: Spring-Verlag, 1995.

[7] VAPNIK VN. An overview of statistical learning theory [J]. IEEE Trans on Neural Network, 1999, 10(11): 988-999.

[8] 郑小霞, 钱锋. 高斯核支持向量机分类和模型参数选择研究[J]. 计算机工程与应用, 2006, 42(1): 77-79.