

分级在线自组织学习的 GD-FNN 算法研究*

左 军¹, 周 灵¹, 孙亚民²

(1. 佛山科学技术学院 电子与信息工程学院, 广东 佛山 528000;
2. 南京理工大学计算机科学与技术学院, 江苏 南京 210094)

摘要: 提出了基于椭圆基函数 (EBF) 的广义动态模糊神经网络 (GD-FNN) 算法。算法提取的模糊规则具有很好可理解性, 可以作为建模工具, 也可以作为知识提取的工具。广义动态模糊神经网络由于基于模糊 ε -完备性, 同时提出了一种新颖的在线参数分配机制, 从而缓解了初始化的随机选择, 且与输入变量不同值域没有关系, 因而更容易构造一个较好性能的模糊系统。开发了仿真程序, 对具体案例进行仿真, 取得了较为理想的结果。

关键词: 广义动态模糊神经网络; 动态模糊神经网络; 径向基函数; 椭圆基函数

中图分类号: TP301.6 **文献标志码:** A **文章编号:** 0529-6579 (2015) 03-0026-05

Research on Online Self-Organizational Learning GD-FNN Algorithm by Grading

ZUO Jun¹, ZHOU Ling¹, SUN Yamin²

(1. School of Electronic and Information Engineering, Foshan University, Foshan 528000, China;
2. School of Computer and Technology, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: General dynamic fuzzy neural network (GD-FNN) algorithm is proposed based on the elliptic basis function (EBF). Fuzzy rules generated from the algorithm are intelligibility. It can be used as a modeling tool. and a tool of knowledge extraction. Because of a novel on-line parameter allocation mechanism for alleviating the random selection in initialization without relation to different input variable range, the proposed GD-FNN based on fuzzy ε -completeness is more easy to construct a good fuzzy system in performance. The simulation program is also developed based on the GD-FNN algorithm and ideal results are achieved by simulation in specific design case.

Key words: generalized dynamic fuzzy neural network (GD-FNN); dynamic fuzzy neural network (D-FNN); radial basis function (RBF); elliptic basis function (EBF)

本文提出了基于椭圆基函数 (EBF) 的广义动态模糊神经网络, 其主要目标是用一种简单、快速的方法来构造一个模糊系统使得: ① 对于知识工程可以获得一些有用的模糊规则。② 可以作为系统建模工具用于控制工程、模式识别等方面。值得指出的是, 尽管 GD-FNN 比 D-FNN 更复杂, 但是, 它所需要的预先设定的参数却比 D-FNN 少, 即使

用户对系统的了解要求更少, 因此, 它比 D-FNN 实现起来更加容易^[1]。

1 GD-FNN 学习算法理论分析

1.1 模糊规则产生准则

定义1 模糊规则的 ε -完备性: 对某个变化范围内的输入, 如果至少存在一条模糊规则, 使得匹

* 收稿日期: 2014-08-24

基金项目: 广东省自然科学基金资助项目 (S2011020002719)

作者简介: 左军 (1958年生), 男; 研究方向: 智能计算、信息安全、数据库应用; E-mail: 13392209106@126.com

配度 (即触发权) 的值不小于 ε , 我们就说这样的模糊系统具备 ε -完备性^[2]。

在模糊系统中, ε 的最小值通常选择为 $\varepsilon_{\min} = 0.5$ 。

从模糊规则观点来看, 一条模糊规则是在输入-输出空间中的一个局部表达。如果一个新的样本满足 ε -完备性, 那么 GD-FNN 不会产生一个新的规则, 而会通过更新现有规则的参数来包容该样本。

根据 ε -完备性, 当一个观测数据 $(X_k, t_k), k = 1, 2, \dots, n$ 进入系统, 可以计算出马氏距离 $md_k(j)$, 它是观测数据 X_k 和现有的 RBF 单元的中心 $C_j(j = 1, 2, \dots, u)$ 之间的马氏距离。

找到

$$J = \arg \min_{1 \leq j \leq u} (md_k(j)) \quad (1)$$

如果

$$md_{k,\min} = md_k(J) > k_d \quad (2)$$

其中, k_d 是一个预先设定并与 ε 相关联的阈值, 则表明现有的系统不满足 ε -完备性而应该考虑产生一条新规则。这里, k_d 在学习的过程中是按下式变化的:

$$k_d = \begin{cases} d_{\max} = \sqrt{\ln(1/\varepsilon_{\min})} & 1 < k < n/3 \\ \max[d_{\max} \times \gamma^k, d_{\min}] & n/3 \leq k \leq 2n/3 \\ d_{\min} = \sqrt{\ln(1/\varepsilon_{\max})} & 2n/3 < k \leq n \end{cases} \quad (3)$$

上式中, k 是学习的次数, $\gamma \in (0, 1)$ 称为衰减常数, 它可以由式 (3) 推算出来:

$$\gamma = \left(\frac{d_{\min}}{d_{\max}} \right)^{3/n} = \left(\frac{\sqrt{\ln(1/\varepsilon_{\max})}}{\sqrt{\ln(1/\varepsilon_{\min})}} \right)^{3/n} \quad (4)$$

1.2 系统误差

GD-FNN 输出误差是确定是否增加新的模糊规则的一个重要因素。误差准则可以表示如下:

对每个观测数据 $(X_k, t_k), k = 1, 2, \dots, n$, 其中, n 是全部训练数据的数量, X_k 是第 k 个输入样本向量, t_k 是第 k 个期望的输出, y_k 为 GD-FNN 的输出。定义系统误差为^[3]:

$$\|e_k\| = \|t_k - y_k\| \quad (5)$$

如果

$$\|e_k\| > k_e \quad (6)$$

则应该考虑增加一条新的模糊规则。这里, k_e 是预先定义的一个阈值, 它在学习过程中按以下准则逐渐变化:

$$k_e = \begin{cases} e_{\max} & 1 < k < n/3 \\ \max[e_{\max} \times \beta^k, e_{\min}] & n/3 \leq k \leq 2n/3 \\ e_{\min} & 2n/3 < k \leq n \end{cases} \quad (7)$$

其中, e_{\min} 是 GD-FNN 期望的输出精度, e_{\max} 是选择的最大误差, k 是学习的次数, $\beta \in (0, 1)$ 称为收敛常数, 可以很容易推导出:

$$\beta = \left(\frac{e_{\min}}{e_{\max}} \right)^{3/n} \quad (8)$$

1.3 高斯宽度修正

在前面的讨论中, 我们提到, 在 $\|e_k\| > k_e$ 和 $md_{k,\min} > k_d$ 的情况下一条新的模糊规则应该产生。但还需要考虑另一种重要情况: $\|e_k\| > k_e$ 和 $md_{k,\min} \leq k_d$, 因为这种情况下的性能不好。这个情况表明, 虽然 X_k 能被相邻的模糊规则包含, 但该规则并没有重要到该 EBF 单元能很好地覆盖数据 X_k , 或者反映系统的输出在这个椭球区域有急剧变化。因此, 该椭球区域应该修正以获得更好的局部逼近^[5]。

对于一个新的样本 X_k , 可以找到以马氏距离“最靠近”该样本的第 j 个规则。如果条件 $\|e_k\| > k_e$ 和 $md_{k,\min} \leq k_d$ 都成立, 把向量 X_k 分解为相应的一维变量, 则输入变量 $x_i (i = 1, 2, \dots, r)$ 的最近的隶属函数的宽度 $\sigma_{ij} (j = 1, 2, \dots, u)$ 可以修正如下:

$$\sigma_{ij}^{\text{new}} = \zeta \times \sigma_{ij}^{\text{old}} \quad (9)$$

其中, $\zeta \in (0, 1)$ 是衰减因子, 它是由如下基于输入变量敏感性的式子决定的:

$$\zeta = \begin{cases} \frac{1}{1 + k_w (B_{ij} - 1/r)^2} & B_{ij} < 1/r \\ 1 & B_{ij} \geq 1/r \end{cases} \quad (10)$$

其中, r 是输入变量的总数。式 (10) 表明, 如果第 i 个输入变量的 ERR 小于与第 j 个规则中输入变量有关的所有项的平均水平, 则第 i 个输入变量的宽度将减小。如果当 $B_{ij} = 0$ 时设 $\zeta_{\min} = k_s$, 则参数 k_w 的值可以由式 (10) 得到, 且 ζ 的最终值为:

$$\zeta = \begin{cases} \frac{k_s}{k_s + r^2 (1 - k_s) (B_{ij} - 1/r)^2} & B_{ij} < 1/r \\ 1 & B_{ij} \geq 1/r \end{cases} \quad (11)$$

1.4 结果参数确定

设 r 个输入变量的 n 次观测样本产生 u 个模糊规则, 把式 $y(x_1, x_2, \dots, x_r) = \sum_{j=1}^u w_j \cdot \varphi_j$ 写成矩阵式为

$$\mathbf{W}\boldsymbol{\varphi} = \mathbf{Y} \quad (12)$$

其中, $\mathbf{W} \in \mathcal{R}^{u(r+1)}$, $\boldsymbol{\varphi} \in \mathcal{R}^{u(r+1) \times n}$ 和 $\mathbf{Y} \in \mathcal{R}^n$ 。假设期望的输出为 $\mathbf{T} = (t_1, t_2, \dots, t_n) \in \mathcal{R}^n$ 。确定最优参数 \mathbf{W}^* 的问题可以用公式表示为最小化 $\|\mathbf{W}\boldsymbol{\varphi} - \mathbf{T}\|_2$ 的线性问题, \mathbf{W}^* 由广义逆方法确定如下

$$\mathbf{W}^* = \mathbf{T}(\boldsymbol{\varphi}^T \boldsymbol{\varphi})^{-1} \boldsymbol{\varphi}^T \quad (13)$$

其中, $\boldsymbol{\varphi}^T$ 是 $\boldsymbol{\varphi}$ 的转置, $\boldsymbol{\varphi}^+ = (\boldsymbol{\varphi}^T \boldsymbol{\varphi})^{-1} \boldsymbol{\varphi}^T$ 是 $\boldsymbol{\varphi}$ 的广义逆^[6]。

2 仿真结果与分析

本文将通过非线性函数逼近以及非线性动态系统辨识实例来验证 GD-FNN 算法的有效性, 并通过与 D-FNN 以及其他方法的比较, 来检验 GD-FNN 在学习效率和性能方面的指标。

例 1: 三输入非线性函数逼近。在这个例子中, 所逼近的三输入非线性函数如下

$$t = (1 + x^{0.5} + y^{-1} + z^{-1.5})^2 \quad (14)$$

该函数被广泛地用来作为检验实例, 如文献 [8] ~ 文献 [11]。训练结果见图 1。

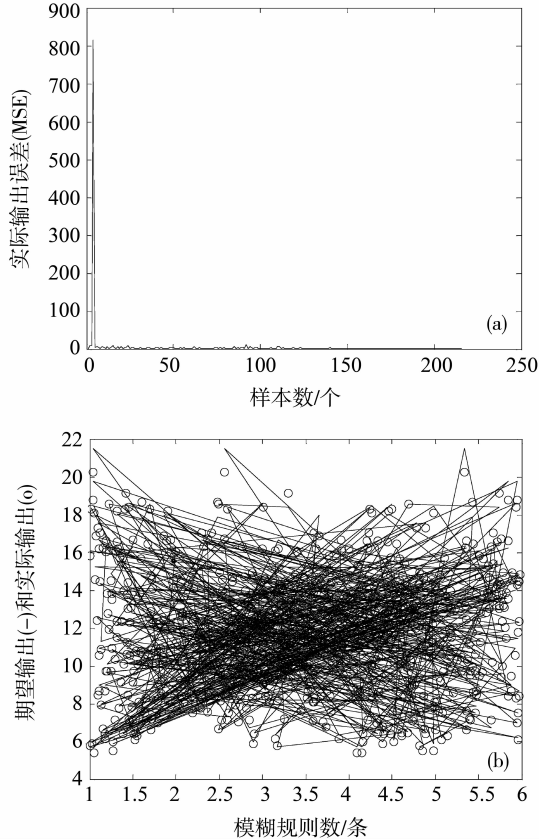


图 1 GD-FNN 用于三输入非线性函数逼近

Fig. 1 GD-FNN for three input nonlinear function approximation

(a) 实际输出误差; (b) 期望 (-) 与实际 (o) 输出

首先, 从输入范围 $[1, 6] \times [1, 6] \times [1, 6]$ 中随机采样到 216 个训练数据。训练结果如图 1 所示。图 1 (a) 是实际输出误差。实际误差很小, 并很快趋于稳定。图 1 (b) 是期望与实际输出的比较图。

为了比较性能, 文献 [11], 文献 [12] 和文献 [13] 采用了相同的性能评判指标。为方便读者了解, 将该性能指标重新列写如下

$$\text{APE} = \frac{1}{n} \sum_{i=1}^n \frac{|t(i) - y(i)|}{|t(i)|} \times 100\% \quad (15)$$

其中, n 是数据对的数目, $t(i)$ 和 $y(i)$ 分别是第 i 个期望输出和实际输出。从相同的工作范围中随机地选择另外 125 个数据, 用来测试学习后的 GD-FNN 的泛化能力。GD-FNN 与 ANFIS、NNDFR、FM 和 OLS 比较见表 1。

表 1 GD-FNN 与其他方法的比较

Table 1 Comparison of GD-FNN with other methods

模型	APE _{tm} /%	APE _{chk} /%	参数 数目	训练集 大小	测试集 大小
GD-FNN	2.11	1.14	64	216	125
ANFIS ^[14]	0.043	1.066	50	216	125
NNDFR ^[10]	3.51	54.9	20	20	20
FM ^[9]	1.5	2.1	22	20	20
OLS ^[15]	2.43	2.56	66	216	125

仿真结果表明, GD-FNN 的训练误差比 ANFIS 大得多, 然而, 它们的测试误差几乎相同。和 NNDFR、FM 和 OLS 方法相比, 在具有同样训练数据和测验数据的条件下, 本文的结果比 NNDFR、FM 和 OLS 好得多。

在此, 我们将通过非线性系统辨识等实例来验证 GD-FNN 算法的有效性, 并与正交最小平方 (OLS)、基于 RBF 适应性模糊系统 (RBF-AFS) 及动态模糊神经网络 (D-FNN) 等做了一些比较。

例 2: 非线性动态系统的辨识。被辨识的对象描述如下

$$y(t+1) = \frac{y(t)y(t-1)[y(t)+2.5]}{1+y^2(t)+y^2(t-1)} + u(t) \quad (16)$$

该对象也被文献 [11], 文献 [12] 及文献 [13] 所采用。该对象给出的无外力系统的平衡状态分别为 (0, 0) 和 (2, 2)。如果串-并行辨识模型用来辨别该对象, 则这个模型可以用如下方程来表示^[16]

$$\hat{y}(t+1) = f(y(t), y(t-1), u(t)) \quad (17)$$

其中, f 是由三输入单输出的 GD-FNN 实现的函数。其中输入 $u(t) = \sin(2\pi t/25)$ 。选择如下参数: $\varepsilon_{\min} = 0.5$, $\varepsilon_{\max} = 0.8$, $e_{\min} = 0.03$, $e_{\max} = 0.5$, $k_{mf} = 0.5$, $k_s = 0.9$ 以及 $k_{err} = 0.003$ 。仿真结果如图 2 所示。

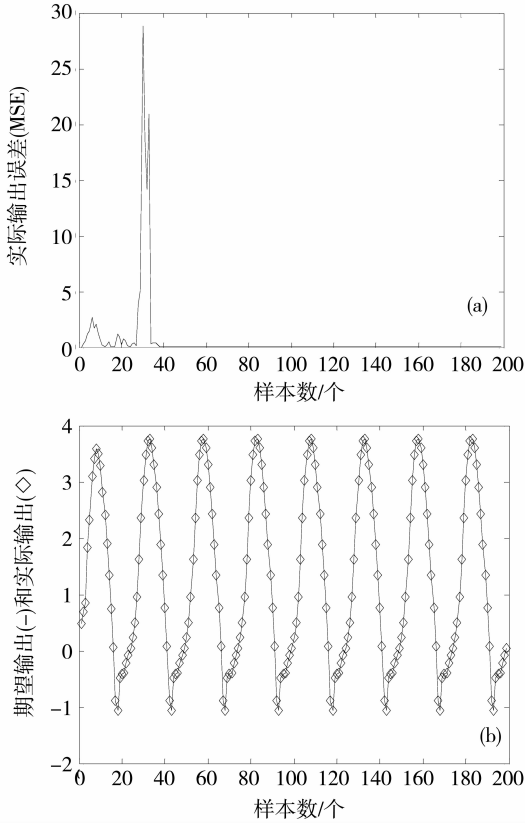


图 2 GD-FNN 在非线性动态系统辨识的应用
Fig. 2 GD-FNN in the application of nonlinear dynamic system identification
(a) 实际输出误差; (b) 辨识结果

从图 2 (a) 中可知, 实际输出误差很小, 说明 GD-FNN 性能优越; 从图 2 (b) 可以看到, GD-FNN 对系统能准确辨识, 其中目标数据用 “-” 表示、辨识数据用 “◇” 表示。

不同方法性能比较见表 2。可以看出 GD-FNN 在一个简单的规则基础上能表现较高的性能。表 2 显示, GD-FNN 的结果比用 BP 训练的 RBF-AFS 的

表 2 不同的方法间的性能比较

Table 2 Comparison of performance between different methods

	GD-FNN	D-FNN	RBF-AFS ^[10]	OLS
规则数	8	6	35	65
参数数目	56	48	280	326
RMSE	0.010 8	0.028 3	0.138 4	0.028 8

结果好得多。由表 2 可以看出接收区域对系统性能起着重要作用, 即使使用更多的模糊规则和可调参数, 使用固定的接收区域 OLS 的性能也不如自适应接收区域的性能。由表 2 可知, 虽然 GD-FNN 和 D-FNN 有相同的规则和参数, 但是 GD-FNN 比 D-FNN 能达到更好的性能。

通过分级的在线自组织学习算法, 我们提出的方法在辨识系统结构的同时修正参数, 而且, 文献 [8] 和文献 [13] 的研究表明, LLS (线性最小二乘) 方法比 BP 算法具有更高效的学习速度。例 1 和例 2 的运行时间列在表 3 中。

表 3 例 1 和例 2 中 GD-FNN 的操作时间
Table 3 Case 1 and case 2 in GD-FNN operation time

序号	样本长度	运行时间/s
1	216	15.4
2	200	12.5

但是, 像 D-FNN, 本方法的一个缺点是, 对于训练样本它不能提供一个全局最优的解。最后的结果会受到预定参数的影响。这一点从表 1 看出。尽管规则数与 ANFIS 几乎相同, 可调节参数比 ANFIS 更多, 但是训练误差却比 ANFIS 大得多。这是为实现快速学习所付出的代价。不过值得一提的是类似的问题也存在在于 BP 算法中, 譬如学习速度和初始化参数选择等。如果仔细地选择预定参数, 则可以达到比 BP 算法更好的性能, 这可以从表 2 看出来。表 2 显示, GD-FNN 的结果比 BP 算法训练的 RBF-AFS 的结果好得多。

3 结 论

在 D-FNN 算法基础上, 提出了一种既适合模糊规则抽取也可用于系统建模的广义动态模糊神经网络 (GD-FNN) 方法。该方法基于椭圆基函数 (EBF), 并提出模糊 ε -完备性作为在线参数分配机制, 避免初始化选择的随机性; 同时, 该算法不仅能对模糊规则而且能对输入变量的重要性作出评价, 从而使每条规则的输入变量的宽度可以根据它对系统性能贡献大小实施在线自适应调整。

参考文献:

[1] 伍世虔, 徐军. 动态模糊神经网络: 设计与应用 [M]. 北京: 清华大学出版社, 2008.
[2] 徐丽娜. 神经网络控制 [M]. 北京: 电子工业出版社, 2009.

- 202.
- [8] 符华·鲍洛金. 弹性体系的动力稳定性[M]. 北京: 高等教育出版社, 1960.
- [9] CHEN J S, YANG C H. Experiment and theory on the nonlinear vibration of a shallow arch under harmonic excitation at the end [J]. *ASME Journal of Applied Mechanics*, 2005, 74 (6): 1061 – 1070.
- [10] 王连华, 易壮鹏, 张辉. 周期荷载作用下几何缺陷拱的动力稳定性 [J]. *湖南大学学报*, 2007, 34: 16 – 19.
- [11] 赵洪金, 刘超. 考虑剪切变形的圆弧深拱参数共振稳定性分析 [J]. *振动与冲击*, 2012, 31: 119 – 122.
- [12] IRIE T, YAMADA G. Natural frequencies of out-of-plane vibration of arches [J]. *Journal of Applied Mechanics*, 1982, 49: 910 – 913.
- [13] SILVA J M M E, URGUEIRA A P V. Out-of-plane dynamic response of curved beams—An analytical model [J]. *International Journal of Solids and Structures*, 1988, 2: 271 – 284.
- [14] KAWAKAMI M, SAKIYAMA T, MATSUDA H, et al. In-plane and out-of-plane free vibrations of curved beams with variable sections [J]. *Journal of Sound and Vibration*, 1995, 187: 381 – 401.
- [15] HUANG C S, TSENG Y P. Out-of-plane dynamic responses of non-circular curved beams by numerical Laplace transform [J]. *Journal of Sound and Vibration*, 1998, 215 (3): 407 – 424.
- [16] LEE S Y, CHAO J C. Out-of-plane vibrations of curved non-uniform beams of constant radius [J]. *Journal of Sound and Vibration*, 2000, 238 (3): 443 – 458.
- [17] 董宁娟, 赵洪金. 开口薄壁截面圆弧拱空间动力稳定性分析 [J]. *科学技术与工程*, 2012, 12: 8569 – 8573.
- [18] 董宁娟, 赵洪金. 闭口薄壁截面圆弧拱空间动力稳定性分析 [J]. *噪声与振动控制*, 2013, 2: 23 – 26.
- [19] R·克拉夫, J·彭津. 结构动力学[M]. 2版. 北京: 高等教育出版社, 2006.
- [20] 卢汉文. 基于 Mathieu-Hill 系统的拱结构动力稳定研究[D]. 广州: 广州大学, 2015.

(上接第 29 页)

- [3] LEONTARITIS I J, BILLINGS S A. Input-output parametric models for nonlinear systems, Part 1: Deterministic nonlinear systems [J]. *Int J Contr*, 2009, 49 (2): 303 – 344.
- [4] 任爱红. 模糊随机过程函数列均方差一致 Henstock 积分的可积性 [J]. *中山大学学报: 自然科学版*, 2010, 51 (4): 41 – 44.
- [5] HONG Z. Algebraic feature extraction of image for recognition [J]. *Patt Recog*, 2007, 24 (2): 211 – 219.
- [6] 杨文光. 权值直接确定的三角形模糊前向神经网络 [J]. *中山大学学报: 自然科学版*, 2013, 52 (2): 33 – 37.
- [7] BRUNELLI R, POGGIO T. Face recognition: features versus templates [J]. *IEEE Trans Pattern Analysis and Machine Intelligence*, 2013, 35 (10): 1042 – 1053.
- [8] LEE C C. Fuzzy logic in controller part I, II [J]. *IEEE Trans Syst Man and Cybern*, 2010, 40: 404 – 436.
- [9] SUGENO M, KANG G T. Structure identification of fuzzy model [J]. *Fuzzy Sets and Systems*, 2008, 38: 15 – 23.
- [10] TAKAGI T, SUGENO M. NN-driver fuzzy reasoning [J]. *Int J Approximate Reasoning*, 2011, 15: 191 – 212.
- [11] CHO K B, WANG B H. Radial basis function based adaptive fuzzy systems and their applications to system identification and prediction [J]. *Fuzzy Sets and Systems*, 2012, 83: 325 – 339.
- [12] CHAO C T, CHEN Y J, Teng C C. Simplification of fuzzy neural systems using similarity analysis [J]. *IEEE Trans Syst, Man, Cybern, Part B: Cybern*, 2013, 35 (2): 344 – 354.
- [13] WANG L X. Adaptive fuzzy systems and control: Design and stability analysis [M]. Prentice-Hall: Englewood Cliffs, NJ, 2011.
- [14] JANG J S R. ANFIS: Adaptive-network-based fuzzy inference system [J]. *IEEE Trans Syst, Man, Cybern*, 2013, 33 (3): 665–684.
- [15] CHEN S, COWAN C F N, GRANT P M. Orthogonal least squares algorithm for radial basis function network [J]. *IEEE Trans Neural Networks*, 2011, 20: 302 – 310.
- [16] 尹华, 胡玉平. 基于随机森林的不平衡特征选择算法 [J]. *中山大学学报: 自然科学版*, 2014, 53 (5): 59 – 65.