

基于非线性动态系统辨识的 D-FNN 算法研究*

杨文茵¹, 张德丰¹, 王传胜²

(1. 佛山科学技术学院计算机系, 广东 佛山 528000;
2. 暨南大学计算机科学系, 广东 广州 510000)

摘要: D-FNN 的基本思想是构造一个基于扩展的 RBF 神经网络, 它可以看成是一个 TSK 模糊系统, 也可以看作是归一化的高斯 RBF 神经网络。D-FNN 算法中, 不仅参数可以在学习过程中调整, 同时, 也可以自动确定模糊神经网络的结构。非线性参数是由训练样本和高斯宽度直接决定的, 只需一步训练就可以达到目标。由于修剪策略的应用, 网络的结构不会持续增长, 因而确保了系统的泛化能力。使用 D-FNN 对非线性动态系统辨识进行了仿真, 并与相关算法作比较, 从而发现了 D-FNN 算法的有效性和高效性。

关键词: 动态模糊神经网络; 模糊规则; 系统辨识; RBF

中图分类号: TP393 文献标志码: A 文章编号: 0529-6579(2014)05-0020-05

Research Based on D-FNN Algorithm on the Nonlinear Dynamic System Identification

YANG Wenyin¹, ZHANG Defeng¹, WANG Chuansheng²

(1. Department of Computer Science, Foshan University, Foshan 528000, China;
2. Department of Computer Science, Jinan University, Guangzhou 510000, China)

Abstract: Dynamic Fuzzy Neural Network (D-FNN), which basic idea is to construct a RBF neural network based on extension, could be seen as a TSK fuzzy system, as well as a Gaussian RBF neural network based on normalized. Within D-FNN algorithms, not only parameters could be adjusted in the learning process, but also the structure of fuzzy neural network could be automatically determined. Nonlinear parameters are directly decided by the training samples and Gaussian width, which only need one step training to achieve this goal. Due to the application of pruning strategies, network structure would not continue to grow, thus ensuring the generalization capability of the system. Simulations are performed on nonlinear dynamic system identification by using D-FNN, and the effectiveness and efficiency of D-FNN algorithm are proved by comparison with related algorithms.

Key words: D-FNN; fuzzy rule; system identification; Radial Basis Function

本文将模糊逻辑和神经网络相结合, 利用神经网络实现模糊化、模糊规则表达、模糊逻辑推理和反模糊化过程, 提出了自适应学习方法, 克服了基于先验知识确定模糊规则及网络结构的困难^[1]。模糊逻辑、神经网络和模糊神经网络是智能计算这门交叉科学的基本内容, 近年来他们都得到了飞速的发展并获得了广泛的应用。多学科算法的相互交

叉、相互渗透可以进一步促进现代科学技术的发展^[2]。

1 RBF 神经网络的学习算法

1.1 模糊神经网络的结构

D-FNN (动态模糊神经网络) 的结构是基于扩展的 RBF (径向基) 神经网络。RBF 神经网络

* 收稿日期: 2014-01-24

基金项目: 广东省自然科学基金资助项目 (S2011020002719)

作者简介: 杨文茵 (1982 年生), 女; 研究方向: 智能计算; E-mail: yangwenyin1982@163.com

具有单隐层的三层前馈网络，RBF 结构如图 1 所示。由于它模拟人脑中局部调整、相互覆盖接收域的神经网络，RBF 网络是一种局部逼近网络，已经证明它能以任意精度逼近任一连续函数^[3]。

设网络输入 n 维向量 \mathbf{u} ，输出 m 维向量 \mathbf{y} ，输入/输出样本对长度为 L 。

RBF 网络隐层第 i 个节点的输出为

$$q_i = R(\|\mathbf{u} - \mathbf{c}_i\|) \quad (1)$$

式中：

\mathbf{u} 为 n 维输入向量； \mathbf{c}_i 为第 i 个隐节点的中心， $i = 1, 2, \dots, s$ ； $\|\cdot\|$ 通常为欧氏范数； $R(\cdot)$ 为 RBF 函数，具有局部感受的特性。它有多种形式，体现了 RBF 网络非线性映射能力^[4]。

网络输出层第 k 个节点的输出为隐节点输出的线性组合，即

$$y_k = \sum_i w_{ki} q_i - \theta_k \quad (2)$$

式中， w_{ki} 为 $q_i \rightarrow y_k$ 的连接权； θ_k 为第 k 个输出节点的阈值^[5]。

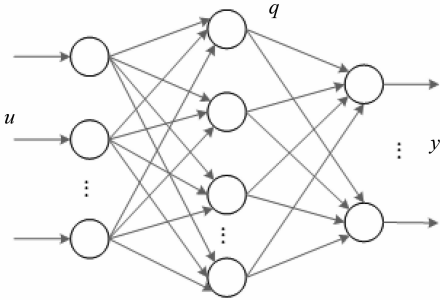


图 1 RBF 网络结构

Fig. 1 RBF network structure

设有 p 组输入/输出样本 \mathbf{u}_p/d_p ， $p = 1, 2, \dots, L$ 定义目标函数（ L_2 范数）为

$$J = \sum_p \|\mathbf{d}_p - \mathbf{y}_p\|_2 = \frac{1}{2} \sum_p \sum_k (d_{kp} - y_{kp})^2 \quad (3)$$

学习的目的是使

$$J \leq \varepsilon \quad (4)$$

式中， \mathbf{y}_p 是在 \mathbf{u}_p 输入下网络的输出向量^[6]。

高斯径向基函数网络隐节点的输出为

$$q_i = R\left[\frac{(\mathbf{u} - \mathbf{c}_i)^T \sum_i^{-1} (\mathbf{u} - \mathbf{c}_i)}{\sigma_i}\right] \quad (5)$$

式中， $[\cdot]$ 表示向量 \mathbf{u} 与 \mathbf{c}_i 间的马氏距离。

当 \sum_i 为对角阵时，式 (5) 为

$$q_i = R\left[\sum_{j=1}^n \frac{(\mathbf{u}_j - \mathbf{c}_{ij})^2}{\sigma_{ij}^2}\right] \quad (6)$$

当 \sum_i 为单位阵时，式 (5) 为

$$q_i = R\left[\frac{(\mathbf{u} - \mathbf{c}_i)^T (\mathbf{u} - \mathbf{c}_i)}{\sigma_i^2}\right] \quad (7)$$

作用函数是高斯径向基函数（RBF），是非线性的，即

$$R(h) = \exp(-h) \quad (8)$$

$$h = \frac{(\mathbf{u} - \mathbf{c}_i)^T (\mathbf{u} - \mathbf{c}_i)}{\sigma_i^2} \quad (9)$$

式中， \mathbf{u} 为 n 维输入向量， $j = 1 \sim n$ ； \mathbf{c}_i 为第 i 个隐节点的中心； σ_i 为第 i 个隐节点的标准化参数； σ_{ij} 为第 i 个隐节点第 j 分量的标准化参数。RBF 神经网络被广泛用于函数逼近以及模式识别中，在这些应用中，样本的维数通常很小^[7]。正如 Moody 和 Darken 在文献 [8] 中所指出的，“RBF 神经网络是最适用于学习逼近连续的或分段连续的实值映射，其中该映射的输入维数充分的小”。

本文提出了一种动态模糊神经网络 D-FNN 结构及其学习算法，该模糊神经网络的结构基于扩展的径向基神经网络。其学习算法的最大特点是参数的调整和结构的辨识同时进行，且学习速度快，可用于实时建模与控制。动态模糊神经网络的结构如图 2 所示，共 5 层。在图 2 中， x_1, x_2, \dots, x_r 是输入的语言变量， y 是系统的输出， MF_{ij} 是第 i 个输入变量的第 j 个隶属函数， R_j 表示第 j 条模糊规则， N_j 是第 j 个归一化节点， w_j 是第 j 个规则的结果参数或者连接权， u 指系统总的规则数。

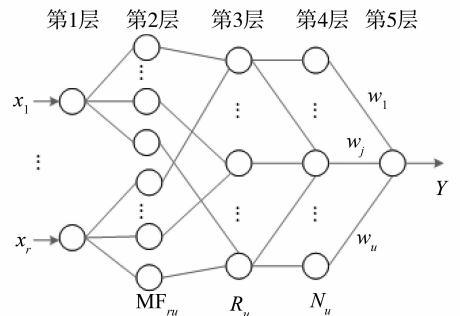


图 2 D-FNN 结构

Fig. 2 D-FNN network structure

1.2 神经网络用于系统辨识

如果实际系统的输出只是依赖于其当前时刻的输入，而与历史输入无关，就称这个系统是静态系统。如果系统的输出不仅依赖于以前的输出而且依赖于当前的输入，则称这类系统为动态系统。 n 阶动态系统通常可以用如下方程描述^[9]：

$$y(t) = f[y^{(n)}(t), \dots, y'(t), x(t)] \quad (10)$$

其中， $y(t)$ 和 $x(t)$ 分别为对象的输出和输入信号，

$y^{(n)}(t)$ 为 $y(t)$ 的 n 阶微分, 它是 $y(t)$ 的以前信息。如果函数 $f(\cdot)$ 是非线性的, 则系统 (10) 称为非线性动态系统。本文将研究这一类型的系统。

文献 [10] 指出, 在较弱的假设下, 任何一个非线性离散时间系统都可用如下的 NARMAX (Nonlinear AutoRegressive Moving Average with eXogenous inputs) 模型来描述:

$$y(t) = f[y(t-1), \dots, y(t-n_y), x(t-1), \dots, x(t-n_x), e(t-1), \dots, e(t-n_e)] + e(t) \quad (11)$$

其中 $y(t)$ 、 $x(t)$ 、 $e(t)$ 分别是系统的输出、系统的输入和噪声, f 是待辨识的未知函数, n_y 、 n_x 、 n_e 分别是输出、输入和噪声的最大延迟, $e(t)$ 是独立的零均值序列。作为 NARMAX 的一个特例, 实际中经常用到如下的 NARX (Nonlinear AutoRegressive with eXogenous inputs) 模型:

$$y(t) = f[y(t-1), \dots, y(t-n_y), x(t-1), \dots, x(t-n_x)] + e(t) \quad (12)$$

基于 NARX 模型, 文献 [9] 对此作了进一步简化, 并提出如下几个没有干扰情况下的模型。

模型 I

$$y(t) = \sum_i^{n_y} \alpha_i y(t-i) + g[x(t-1), \dots, x(t-n_x)] \quad (13)$$

模型 II

$$y(t) = f[y(t-1), \dots, y(t-n_y)] + \sum_{i=1}^{n_x} \beta_i x(t-i) \quad (14)$$

模型 III

$$y(t) = f[y(t-1), \dots, y(t-n_y)] + g[x(t-1), \dots, x(t-n_x)] \quad (15)$$

基于上述模型, 文献 [9] 提出了如何利用神经网络来进行系统辨识。

用神经网络来进行系统辨识, 一个显而易见的方法就是使神经网络或模糊神经网络的输入-输出结构选择和被辨识的系统一模一样。在实际操作上, 可以有如下两种模型。

1) 并行模型

$$y(t) = \hat{f}[\hat{y}(t-1), \dots, \hat{y}(t-n_y), x(t-1), \dots, x(t-n_x)] \quad (16)$$

2) 串-并行模型

$$\hat{y}(t) = \hat{f}[y(t-1), \dots,$$

$$y(t-n_y), x(t-1), \dots, x(t-n_x)] \quad (17)$$

其中, \hat{f} 是一个神经网络或模糊神经网络, y 、 \hat{y} 分别是待辨识模型的输出和系统的实际输出。

2 实验结果与分析

在并行模型中, 由于反馈模型输入 $\hat{y}(t-i)$ 依赖辨识模型的参数, 因此, 线性参数模型的参数求解变成了一个非线性问题。一般地说, 非线性方法求解优化问题的速度非常慢。因此串-并行方法被研究者广泛使用, 特别是考虑到稳定性问题, 该方法显得非常有优势, 因为它不涉及反馈问题。然而当处理有噪声的系统时, 并行模型应该是首选的, 因为它避免了由于实际系统输出的噪声而导致的偏置问题。本文仿真在 D-FNN 中只使用了串-并行方法。

被辨识的对象描述如下:

$$y(t+1) = \frac{y(t)y(t-1)[y(t)+2.5]}{1+y^2(t)+y^2(t-1)} + u(t) \quad (18)$$

这个模型广为研究人员使用, 比如, 文献 [11] 和文献 [12], 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识该对象, 则这个模型可以用如下方程来表示:

$$\hat{y}(t+1) = f(y(t), y(t-1), u(t)) \quad (19)$$

其中, f 是三输入单输出的 D-FNN。使用同文献 [13] 和文献 [14] 相同的输入 $u(t) = \sin(2\pi t/25)$, 其结果如图 3 所示。其中图 3 (a) 表示产生的模糊规则, 图 3 (b) 表示训练时的实际输出误差, 图 3 (c) 表示均方根误差。与文献 [11]、文献 [12] 及文献 [13] 结果比较如表 1 所示。

由表 1 可以看到, 要达到相同的训练误差, 只有结构自适应的 OLS 方法需要 65 条规则, 远大于 D-FNN 所需要的规则数。文献 [11] 和文献 [12] 也采用了修剪计算^[15], 其所用的规则数虽然比 OLS 方法少, 但还是比 D-FNN 的方法多多了, 而且训练误差比 D-FNN 大得多。由图 3 (d) 可以看到, 尽管 D-FNN 的结构很小, 但其泛化能力很强。结果表明: 与其它算法相比, D-FNN 无论在结构上还是性能上都具有明显的优势。

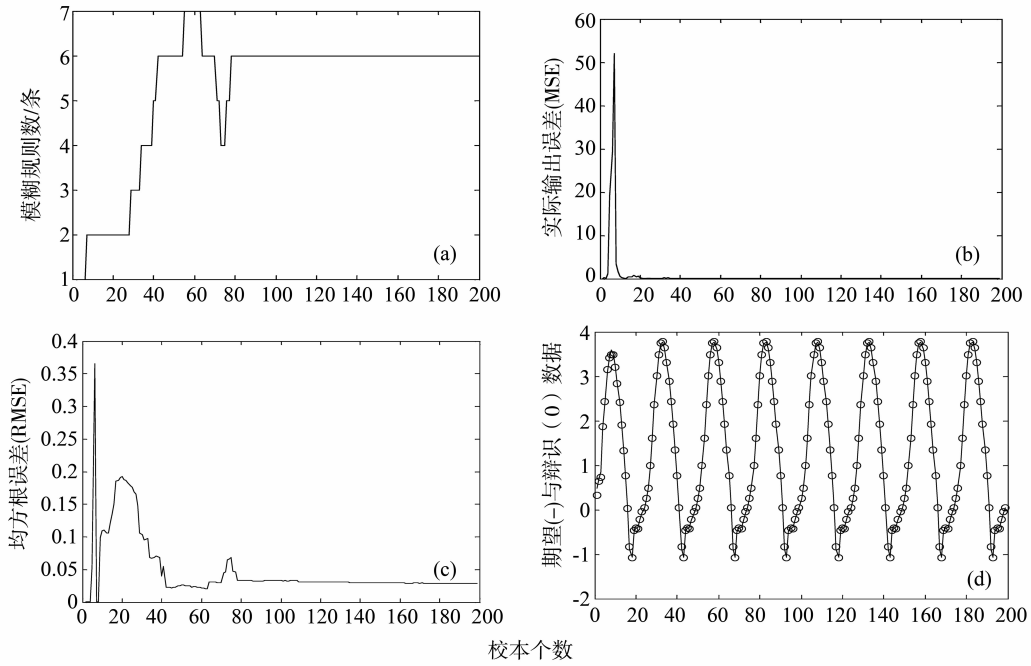


图 3 非动态线性系统的辨识

Fig. 3 Identification of non-linear dynamic systems

(a) 产生的模糊规则; (b) 训练时实际输出误差; (c) 均方根误差; (d) 期望 (-) 与辨识 (o) 数据

表 1 不同算法的性能和结构比较

Table 1 Comparison of performance and structure with different algorithms

	D-FNN	OLS ^[13]	RBF-AFS ^[11]	FNS ^[12]
规则数	7	65	35	22
参数数	47	326	280	84
均方根误差	0.028 2	0.028 8	0.138 4	未列出

为了检验 D-FNN 的泛化能力, 设 $p = 6, n = 500$ 。其中 p 为前 6 步预测结果, n 为 500 个样本对。选择不同参数, 将得到两种不同的神经网络结构, 其训练和测试结构列于表 2 中。

表 2 $n = 500, p = 6$ 时训练数据的仿真结构

Table 2 Simulation structure of training data when $n = 500, p = 6$

	规则数	训练的 RMSE	测试的 REMS
结构 1	5	0.013 1	0.013 4
结构 2	10	0.008 3	0.008 4

当 p 等于 6, 模糊规则数为 10 时, 图 4 说明了 D-FNN 泛化性的测试, 其中, 前 500 个点用于训练, 后 500 个点用于预测。由图 4 及表 2 可以看出, 无论是采样 5 条规则还是 10 条规范, 所构建的模糊神经网络都具有很强的泛化能力, 因为测试误差几乎等同训练误差^[16]。表 3 列出了 D-FNN 与

ANFIS (基于自适应神经网络的模糊推理系统)、OLS (正交最小二乘网络) 和 RBF-AFS (径向基自适应模糊系统) 的比较结果。

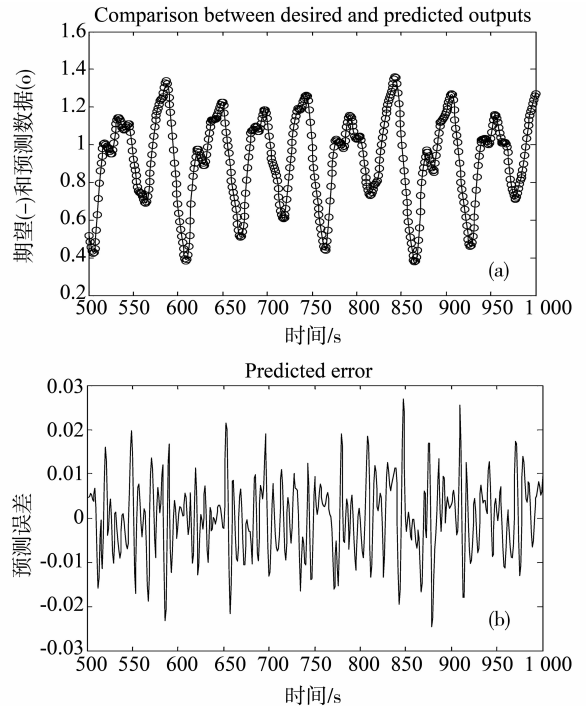


图 4 D-FNN 预测的测试结果

Fig. 4 Test result of D-FNN prediction
(a) 期望 (-) 和辨识 (o) 数据; (b) 预测误差

表 3 不同的结构和性能比较 (训练条件: $n=500, p=6$)Table 3 Structure and performance comparison (Training condition: $n=500, p=6$)

方法	规则数	参数的数量	训练的均方根误差	测试的均方根误差
D-FNN	10	100	0.008 2	0.008 3
ANFIS	16	104	0.001 6	0.001 5
OLS	35	211	0.008 7	0.008 9
RBF-AFS	21	210	0.010 7	0.012 8

比较结果显示,即使 D-FNN 有更多的可调参数,它的性能也并没有 ANFIS 好。原因是 ANFIS 通过迭代学习的方式训练,从而可以达到整体的最优解。而 D-FNN 只能获得次优的结果。然而,与 RAN、RANEKF 及 M-RAN (它们也只能得到次优解)相比, D-FNN 结构小而泛化能力更强。

3 结 论

D-FNN 算法中,非线性参数是由训练样本和启发方法直接决定的,而没有用优化算法来确定,虽然高斯宽度在学习时可以自适应地调整,但学习规则却很简单。D-FNN 算法中,由于分级学习策略的使用可以缓解学习过程中的震荡问题;修剪技术的应用,使得网络的结构不会持续增长。为检验 D-FNN 的有效性,把 D-FNN 与其它的学习算法进行了比较,并深入探讨这些算法与其它算法的相互关系,研究表明 D-FNN 具有简洁的结构和优良的性能。

参考文献:

- [1] JANG J-S R. ANFIS: Adaptive-network-based fuzzy inference system [J]. IEEE Transactions on Systems, Man, and Cybernetics, 2013, 23(3): 665-684.
- [2] KADIRKAMANATHAN V, NIRANJAN M. A function estimation approach to sequential learning with neural networks[J]. neural computation, 2012, 16(4): 954-975.
- [3] HONG Z Q. Algebraic feature extraction of image for recognition[J]. Pattern Recognition, 2011, 24(3): 211-219.
- [4] 徐丽娜. 神经网络控制[M]. 北京: 电子工业出版社, 2009.
- [5] WU S Q, ER M J. Dynamic fuzzy neural networks: a novel approach to function approximation[J]. IEEE Transactions on Systems, Man, and Cybernetics; Part B, 2000, 30(2): 358-364.
- [6] ER M J, WU S Q. A fast learning algorithm for parsimonious fuzzy neural systems[J]. Fuzzy Sets and Systems, 2002, 126: 337-351.
- [7] 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.
- [8] MOODY J, DARKEN C J. fast leaning in network of locally-tuned processing units [J]. Neural Computation, 2009, 1: 281-294.
- [9] NARENDRA K S, PARTHASARATHY K. Identification and control of dynamical system using neural networks [J]. IEEE Transactions on Neural Networks, 1990, 1(1): 4-27.
- [10] LEONTARITIS I J, BILLINGS S A. Input-output parametric models for nonlinear systems, part 1: deterministic nonlinear systems[J]. International Journal of Control, 2006, 41(2): 303-344.
- [11] CHO K B, CHEN B H. Radial basic fuction based adaptive fuzzy systems and their applications to system identification and prediction[J]. Fuzzy Sets and Systems, 2006, 83(3): 325-339.
- [12] CHAO C T, CHEN Y J, TENG C C. Simplification of fuzzy neural systems using similarity analysis[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B; Cybernetics, 2008, 26(2): 344-354.
- [13] CHEN S, COWAN C N, GRANT P M. Orthogonal least squares learning algorithm for radial basis function network [J]. IEEE Transactions on Neural Networks, 2009, 2(2): 302-309.
- [14] 何正风, 张德丰, 孙亚民. 高斯激活函数特征值分解修剪技术的 D-FNN 算法研究[J]. 中山大学学报: 自然科学版, 2013, 52(1): 34-39.
- [15] 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.