

文章编号: 1001-0920(2004)03-0349-02

有理式多层前馈神经网络

梁久祯¹, 何新贵²

(1. 浙江师范大学 计算机科学研究所, 浙江 金华 321004; 2. 北京大学 信息科学与技术学院, 北京 100084)

摘要: 提出了有理式多层前馈神经网络的数学模型, 给出了有理式多层神经网络的学习算法. 就计算复杂度而言, 有理式神经网络的学习算法与传统的多层神经网络反传播算法是同阶的. 函数逼近应用实例结果表明, 将有理式多层神经网络用于解决传统问题是有效的.

关键词: 有理式; 神经网络; 学习算法; 函数逼近

中图分类号: TP183 **文献标识码:** A

Rational fraction multiplayer feed forward neural networks

L I A N G J i u - z h e n¹, H E X i n - g u i²

(1. Institute of Computer Science, Zhejiang Normal University, Jinhua 321004, China; 2. College of Information Science and Technology, Peking University, Beijing 100084, China. Correspondent: L I A N G J i u - z h e n; E-mail: liangjiuzhen@yahoo.com)

Abstract: A mathematic model of rational fraction multiplayer feed forward neural networks is proposed. A learning algorithm for rational fraction multiplayer neural networks is presented. The learning algorithm has the same degree of computing complexity as traditional multiplayer neural networks. The function approximation is also discussed. Experiment result illustrates the effectiveness of the rational fraction multiplayer feed forward neural networks in solving traditional problems.

Key words: rational fraction; neural networks; learning algorithm; function approximation

1 引言

近年来,人们对多层前馈神经网络模型拓扑结构作了改进,不再是简单的只有相邻层神经元的互连接(如考虑输入层神经元直接与输出层神经元的连接)^[1],从而增大了网络的表达能力,减少了隐层神经元的数目.将神经元按输入层到输出层的顺序级联在一起^[2],网络的映射能力和学习算法的收敛速度都有较大的改进.认知机模型^[3]考虑了感受野神经元分为兴奋型神经元和拟制型神经元,通过多层映射实现对模式的分类,具有较高的容错能力.

对多层神经网络函数逼近能力的研究一直是神经网络函数逼近理论研究的热点,并取得了许多成

果^[4].但是,这些结论并没有明确指出对于一类函数问题用什么样的多层前馈神经网络,以较少的代价实现.具有奇异值函数的逼近问题是多项式逼近理论中遇到的难题,也是神经网络函数逼近中较难实现的问题,即使大量增加隐单元的个数也难以满足函数逼近精度,而且学习时间较长.

2 有理式多层前馈神经网络模型

有理式多层前馈神经网络可以看作是对传统多层前馈神经网络模型和认知机模型的改进和推广.设输入层有 n 个神经元,输入向量为 $X = (x_1, x_2, \dots, x_n)$.隐层神经元可分为分子和分母部分,其中分子和分母部分各有 p 个神经元.输入层到隐层

收稿日期: 2003-02-23; 修回日期: 2003-05-30

作者简介: 梁久祯(1968—),男,山东东平人,副教授,博士,从事智能计算与智能信息处理的研究;何新贵(1938—),男,浙江浦江人,中国工程院院士,研究员,从事智能计算等研究.

分子部分的连接权向量为 $w_i = (w_{i1}, w_{i2}, \dots, w_{ip}), i = 1, 2, \dots, n$; 输入层到隐层分母部分的连接权向量为 $v_i = (v_{i1}, v_{i2}, \dots, v_{ip}), i = 1, 2, \dots, n$. 输出层有 m 个神经元, 输出向量为 $Y = (y_1, y_2, \dots, y_m)$, 隐层到输出层的连接权向量为 $u_j = (u_{j1}, u_{j2}, \dots, u_{jm}), j = 1, 2, \dots, p$. 有理式前馈神经网络的结构如图 1 所示. 图中第 2 层单元实线部分表示分子单元及其连接权, 虚线部分表示分母单元及其连接权; 第 3 层单元表示有理式整合部分, 第 2 层与第 3 层神经元之间的连接不是全连接, 且连接权为固定值 1. 设隐层单元(第 3 层单元)的输出为线性输出, 输出层单元传递函数为 Sigmoid 型函数

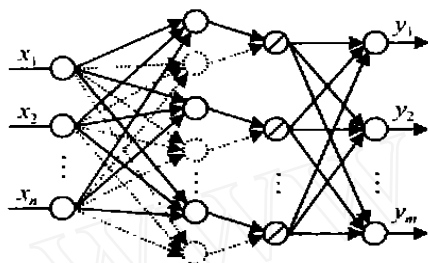


图 1 含一个隐层的有理式前馈神经网络
隐层分子部分单元的输出为

$$h_{uj} = \sum_{i=0}^n w_{ij} x_i \quad (1)$$

隐层分母部分单元的输出为

$$h_{dj} = \sum_{i=0}^n v_{ij} x_i \quad (2)$$

第 3 层单元的整合运算为

$$h_j = h_{uj} / g(h_{dj}). \quad (3)$$

为防止出现零分母, 可取 $g(h_{dj}) = 1 + (h_{dj})^2$. 输出层各单元的输出为

$$y_k = f \left(\sum_{j=0}^p u_{jk} h_j \right), k = 1, 2, \dots, m. \quad (4)$$

3 学习算法

对具有连续传递函数的有理式前馈神经网络, 其学习算法可借鉴传统的前馈神经网络的误差反向传播算法. 定义网络实际输出与期望输出的平方和误差函数为

$$E = \frac{1}{2} \sum_{k=1}^m (y_k - \hat{y}_k)^2. \quad (5)$$

其中 \hat{y}_k 表示期望输出. 由梯度下降算法, 网络权值修正公式如下:

$$w_{ij} = w_{ij} - \alpha \frac{\partial E}{\partial w_{ij}}, \quad (6)$$

$$i = 0, 1, \dots, n, j = 1, 2, \dots, p;$$

$$v_{ij} = v_{ij} - \beta \frac{\partial E}{\partial v_{ij}}, \quad (7)$$

$$i = 0, 1, \dots, n, j = 1, 2, \dots, p;$$

$$u_{jk} = u_{jk} - \lambda \frac{\partial E}{\partial u_{jk}}, \quad (8)$$

$$j = 0, 1, \dots, p, k = 1, 2, \dots, m.$$

式中: α, β 和 λ 为学习步长; 而

$$\frac{\partial E}{\partial v_{ij}} = \frac{x_i}{g(h_{dj})} \sum_{k=1}^m (y_k - \hat{y}_k) f \left(\sum_{j=0}^p u_{jk} h_j \right) u_{jk}; \quad (9)$$

$$\frac{\partial E}{\partial v_{ij}} = - \frac{2h_{dj} h_{uj} x_i}{g(h_{dj})^2} \sum_{k=1}^m (y_k - \hat{y}_k) f \left(\sum_{j=0}^p u_{jk} h_j \right) u_{jk}; \quad (10)$$

$$\frac{\partial E}{\partial u_{jk}} = h_j (y_k - \hat{y}_k) f \left(\sum_{j=0}^p u_{jk} h_j \right). \quad (11)$$

4 应用实例

函数逼近问题. 选取具有奇异值的三角函数作为算例, 计算函数 $f(x) = \sin(1/x)$ 在区间 $[1/(2\pi + 1), 1]$ 上 11 个节点的函数值(即 $f(x) = \sin x$ 在一个周期 $[1, 2\pi + 1]$ 上的等距节点的函数值), 如表 1 所示.

表 1 $f(x) = \sin(x)$ 在 $[1, 2\pi + 1]$ 区间 11 个节点上的函数值

x	$f(x)$
0.137 3	0.841 4
0.150 3	0.363 0
0.165 9	- 0.254 0
0.185 3	- 0.774 0
0.209 7	- 0.998 4
0.241 5	- 0.841 4
0.284 6	- 0.363 1
0.346 6	0.253 9
0.443 1	0.773 9
0.614 1	0.998 3
1.000 0	0.841 5

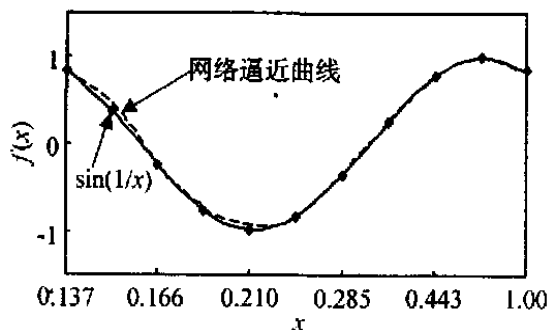


图 2 有理式多层神经网络函数逼近

(下转第 354 页)

各个子 T-S 模糊系统的隶属度函数为

$$\begin{aligned} M_{111}(x_{11}(k)) &= \exp[-2x_{11}^2(k)], \\ M_{211}(x_{11}(k)) &= 1 - M_{111}(x_{11}(k)), \\ M_{112}(x_{12}(k)) &= \frac{1}{1 + \exp[-x_{12}(k)]}, \\ M_{212}(x_{12}(k)) &= 1 - M_{112}(x_{12}(k)), \\ M_{113}(x_{13}(k)) &= \frac{1}{1 + \exp(-2x_{13}(k))}, \\ M_{213}(x_{13}(k)) &= 1 - M_{113}(x_{13}(k)). \end{aligned}$$

利用 Matlab 的 LMI 工具箱中的函数“feasp”求解定理 2 中的 LMI 条件, 可得出使该关联模糊大系统闭环渐近稳定的分散化 PDC 控制器为

$$\begin{aligned} F_{11} &= [-0.0388 \quad -0.4100], \\ F_{21} &= [-0.0376 \quad -0.5142], \\ F_{12} &= [0.4710 \quad 0.3834], \\ F_{22} &= [0.5258 \quad 0.4084], \\ F_{13} &= [-0.0233 \quad -0.0600], \\ F_{23} &= [0.0329 \quad -0.0590]. \end{aligned}$$

4 结 语

本文考虑了离散模糊大系统的分散镇定问题, 整个模糊大系统由 N 个相互关联的子 T-S 模糊系统构成. 基于李亚普诺夫稳定性理论及大系统分散控制理论, 给出了保证该模糊大系统闭环渐近稳定的分散化并行分布补偿(PDC)控制器的设计方法. 这些分散控制器的存在条件均以 LMI 的形式出现, 因此可通过 MATLAB 的 LMI 工具箱有效地对其进行求解. 数值仿真例子说明了本文方法的有效性.

(上接第 350 页)

设计有理式前馈神经网络的拓扑结构为 1-8-1, 即输入层只有 1 个神经元, 隐层有 18 个神经元(分子、分母和整合层均有 6 个神经元), 输出层有 1 个神经元. 经过 42 718 次学习之后, 网络的输出误差为 0.001 003 833 847 214 56, 如图 2 所示.

5 结 论

有理式前馈神经网络是传统前馈神经网络的一般形式, 传统前馈神经网络所具有的性质(如函数逼近和计算能力等)也适用于有理式前馈神经网络. 就计算复杂度而言, 有理式前馈神经网络 BP 算法与前馈神经网络 BP 算法相比是同阶的.

参考文献(References):

- [1] Takagi T, Sugeno M. Fuzzy identification of systems and its application to modeling and control[J]. *IEEE Trans on Systems, Man and Cybernetics*, 1985, 15(1): 116-132.
- [2] Tanaka K, Sugeno M. Stability analysis and design of fuzzy control systems[J]. *Fuzzy Sets Systems*, 1992, 45(2): 136-156.
- [3] Wang H O, Tanaka T, Griffin M F. An approach to fuzzy control of nonlinear systems: Stability and design issues[J]. *IEEE Trans on Fuzzy Systems*, 1996, 4(1): 14-23.
- [4] 陈跃鹏, 张庆灵, 张国峰. 具有对称结构的广义大系统的 H_∞ 分散控制和二次能稳[J]. *东北大学学报(自然科学版)*, 2000, 21(6): 675-677.
(Chen Y P, Zhang Q L, Zhang G F. Decentralized H_∞ control and quadratic stabilization for large-scale descriptor symmetric composite systems [J]. *J of Northeastern University (Natural Science)*, 21(6): 675-677.)
- [5] Yan X G, Dai G Z. decentralized output feedback robust control for nonlinear large-scale systems[J]. *Automatica*, 1998, 34(11): 1469-1472.
- [6] Hsiao F H, Hwang J D. Stability analysis of fuzzy large-scale systems[J]. *IEEE Trans on Systems Man and Cybernetics*, 2001, 32(1): 122-126.
- [7] Boyd S, Ghaoui L, Feron E, et al. *Linear Matrix Inequalities in Systems and Control Theory* [M]. Philadelphia: SIAM, 1994.
- [8] Gahinet P, Nemirovski A, Laub A J, et al. *LMI Control Toolbox* [M]. Natick: Mathworks, 1995.

参考文献(References):

- [1] Shah J V, Poon Chi-Sang. Linear independence of internal representations in multilayer perceptrons[J]. *IEEE Trans on Neural Networks*, 1999, 10(1): 10-18.
- [2] Fahman S E, Lebiere C. The cascade-correlation learning architecture [A]. *Advance in Neural Information Processing System* [C]. San Mateo: Morgan Kaufmann, 1994. 524-532.
- [3] Fukushima K. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position[J]. *Pattern Recognition*, 1982, 15(6): 455-469.
- [4] 梁久祯, 何新贵. 模糊神经网络逼近理论及学习算法研究[D]. 北京: 北京航空航天大学, 2001. 6-11.