

一种基函数神经网络最优隐神经元数目快速确定算法

肖秀春^{1,2,3}, 姜孝华², 张雨浓²

(1 广东海洋大学 信息学院, 广东 湛江 524025; 2 中山大学 信息科学与技术学院, 广东 广州 510275;
3. 浙江大学 CAD/CG 国家重点实验室, 浙江 杭州 310058)

摘 要: 以线性无关的基函数作为隐层神经元的激励函数, 构建了一类基函数神经网络, 且推导出该类神经网络的学习算法; 在此基础上, 设计了一种基于指数增长和折半删减的快速最小隐神经元数目确定算法. 仿真实验表明, 此算法能自适应地、快速有效地确定网络最小隐层神经元数目.

关键词: 广义多项式; 神经网络; 结构自适应确定; 指数增长; 折半删减

中图分类号: TP183

文献标识码: A

文章编号: 1000-7180(2010)01-0057-04

Basis Function Neural Network and Its Fast Optimal-structure Determination

XIAO Xiu-chun^{1,2,3}, JIANG Xiao-hua², ZHANG Yu-nong²

(1 College of Information, Guangdong Ocean University, Zhanjiang 524088, China;

2 School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510275, China;

3 State Key Lab of CAD&CG, Zhejiang University, Hangzhou 310058, China)

Abstract: In this paper, a basis function neural network is constructed, of which the hidden-layer neurons are activated with linear independence basis function. Accordingly, the learning algorithm for the constructed neural network is derived and a fast algorithm based on exponential-growth and binary-delete search strategy is proposed to determinate the optimal number of hidden-layer neurons. The simulation results substantiate that our algorithm can adaptively, quickly and efficiently determine number of hidden neurons in the neural network.

Key words: general polynomial; neural network; structure-adaptive-determination; exponential growth; binary search

1 引言

通常神经网络的性能与其隐层结构具有非常密切的关系^[1], 一个规模较大的隐层结构能很好地学习训练样本, 从而获得较小的输出误差. 但是, 过大规模的隐层结构往往会导致神经网络的过拟合现象, 且其计算时间和空间复杂度较高, 其泛化能力也较差. 若隐层规模过小, 则难以完成对训练样本的学习而达到预定的精度要求. 因此, 神经网络的规模应该适当.

近年来, 人们已经相继提出了一些隐层结构的

设计方法. 根据隐神经元的演变进程不同, 可以把这些方法概括为两类^[2]:

(1) 隐神经元生长

由一个较小规模神经网络结构开始, 逐渐增加隐神经元, 从而增强网络逼近性能, 直到网络的输出误差满足实际问题的精度要求^[3]. 著名的隐神经元生长算法有 Fahlman 等提出瀑流关联等^[4].

(2) 隐神经元删减

在开始时构造一个明显冗余的较大规模的隐层结构, 通过训练使该网络满足实际应用时的精度要求, 然后, 按照某种优选策略删除隐层结构中不重要

收稿日期: 2008-12-31; 修回日期: 2009-03-01

基金项目: 国家自然科学基金项目(60643004, 60775050); 浙江大学 CAD/CG 国家重点实验室开放课题(A0908)

或冗余的隐神经元,以增强神经网络的泛化性能和提高其计算效率^[5-8].

文中以线性无关的基函数作为隐层神经元激励函数,构建了一类基函数神经网络,依据梯度下降法原理,推导出该类神经网络的最优参数迭代修正求解式.然后,针对该类神经网络逼近性能随隐层规模增长而增强的特性,设计了一种基于指数增长和折半删减的快速最小隐神经元数目确定算法.

2 网络拓扑结构

为叙述简便起见,可定义文中神经网络拓扑结构为简单的三层前向结构,即由输入层、隐层和输出层构成,图1所示为单输入单输出的基函数神经网络模型.需要特别指出的是,其隐层的 n 个神经元使用任意一簇线性无关的基函数作为其激励函数,因此,图示神经网络模型是一类神经网络的统一模型.文中将在该神经网络拓扑结构的基础上,推导出其学习算法及其最小隐神经元数目快速、自适应确定算法^[9-10].

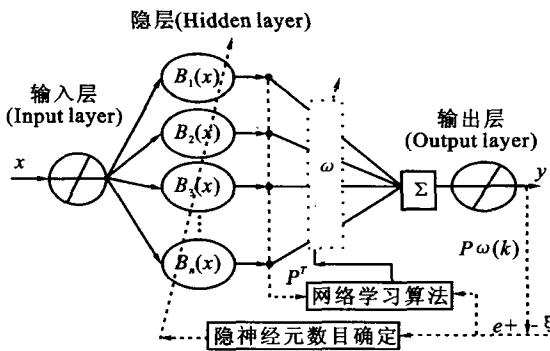


图1 基函数神经网络拓扑结构

3 网络学习算法

对于图1所给出基函数神经网络模型,设其隐神经元数为 n ,隐神经元与输出层神经元之间的权值记为 $\omega_i, i = 1, 2, \dots, n$,隐层神经元激励函数 $B_i(x)$ 为任意线性无关的连续函数^[10].

如果采用训练样本集 $\{(x_j, y_j), j = 1, 2, \dots, m\}$, m 为样本总量,则依据梯度下降法原理,可以定义基函数神经网络的优化目标函数 E 如下:

$$E = \frac{1}{2} \sum_{j=1}^m \left(\sum_{i=1}^n \omega_i B_i(x_j) - y_j \right)^2 \quad (1)$$

于是,有网络参数修正迭代式如下:

$$\omega(k+1) = \omega(k) - \eta \frac{\partial E}{\partial \omega} \bigg|_{\omega=\omega(k)} \quad (2)$$

运用微分学原理,将式(1)中 E 对参数向量 ω

求偏导,有

$$\begin{aligned} \frac{\partial E}{\partial \omega} &= \frac{\partial \left(\frac{1}{2} \sum_{j=1}^m \left(\sum_{i=1}^n \omega_i B_i(x_j) - y_j \right)^2 \right)}{\partial \omega} \\ &= \begin{bmatrix} B_1(x_1) & B_1(x_2) & \cdots & B_1(x_m) \\ B_2(x_1) & B_2(x_2) & \cdots & B_2(x_m) \\ \vdots & \vdots & \ddots & \vdots \\ B_n(x_1) & B_n(x_2) & \cdots & B_n(x_m) \end{bmatrix} \times \\ &\quad \left(\begin{bmatrix} \sum_{j=1}^m \omega_i(k) B_i(x_1) \\ \sum_{j=1}^m \omega_i(k) B_i(x_2) \\ \vdots \\ \sum_{j=1}^m \omega_i(k) B_i(x_m) \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \right) \\ &= \begin{bmatrix} B_1(x_1) & B_1(x_2) & \cdots & B_1(x_m) \\ B_2(x_1) & B_2(x_2) & \cdots & B_2(x_m) \\ \vdots & \vdots & \ddots & \vdots \\ B_n(x_1) & B_n(x_2) & \cdots & B_n(x_m) \end{bmatrix} \times \\ &\quad \left(\begin{bmatrix} B_1(x_1) & B_2(x_1) & \cdots & B_n(x_1) \\ B_1(x_2) & B_2(x_2) & \cdots & B_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ B_1(x_m) & B_2(x_m) & \cdots & B_n(x_m) \end{bmatrix} \right. \\ &\quad \left. \begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_m \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \right) \end{aligned}$$

若令

$$P = \begin{bmatrix} B_1(x_1) & B_1(x_2) & \cdots & B_n(x_1) \\ B_1(x_2) & B_2(x_2) & \cdots & B_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ B_1(x_m) & B_2(x_m) & \cdots & B_n(x_m) \end{bmatrix} \in R^{m \times n}$$

$$\xi = [y_1 \ y_2 \ \cdots \ y_m]^T \in R^m$$

$$\omega = [\omega_1 \ \omega_2 \ \cdots \ \omega_n]^T \in R^n$$

则上式可简写为

$$\frac{\partial E}{\partial \omega} = P^T(P\omega - \xi)$$

将其代入式(2),可得基函数神经网络学习算法为

$$\omega(k+1) = \omega(k) - \eta(P^T(P\omega(k) - \xi))$$

式中, $\eta > 0$ 为学习率, $k = 0, 1, 2, \dots$ 为迭代次数.

4 网络隐神经元数目确定算法

由基函数神经网络模型的特点可知,在能满足给定精度的隐层结构中,隐层神经元数目最小的结构就是最佳结构.因此,文中提出了一种基于指数增长和折半删减搜索策略的隐神经元确定算法.本文设计的自适应确定算法可分两步进行:

(1) 先采用指数式增长搜索把最小神经元数目快速限定在某个范围 $[2^{h-1}, 2^h]$,其中 h 为整数.考虑到神经元个数不能够无限地增长,我们设定隐神经元数最大值为512($h=9$).

(2) 在由(1)限定的范围 $[2^{h-1}, 2^h]$ 内利用二分搜索策略迅速确定最小隐神经元数目.

在后续的仿真工作中将进一步验证,这样设计的算法能自适应地快速确定网络的最小结构.

5 仿真验证

为了验证文算法的可行性及所构造基函数神经网络的有效性,我们进行了大量仿真实验.图2是以目标函数 $y = \cos((5x^2 + 1)(x^2 + 1)/(2 + (x - 1)^3))$ 为例时的仿真结果图,其中图2(a)~(d)分别为设定目标均方误差为 $10^{-1}, 10^{-2}, 10^{-3}, 10^{-25}$ 时的仿真结果,对应由本文算法确定的隐神经元数分别为:6个,122个,243个,299个.另外,由图3(d)可以看出,文中算法能以非常高(实例为 6.84×10^{-26})的精度逼近目标函数.

为进一步了解文中隐神经元数目快速确定算法的可行性与有效性,列出了表1所示的对比数据.表中数据分别为采用文中算法与基于自然数增长方式的隐神经元数确定算法^[2,5]搜索次数的比较.表中数据表明,相比较而言,文中算法能够较大幅度地减少搜索次数,即,算法搜索次数由 $O(n)$ 降低到 $O(\log_2(n))$.

表1 文中算法/自然数增长算法搜索次数对比

目标 MSE 精度	自然数增长法	文中算法
10^{-1}	6	6
10^{-2}	122	14
10^{-3}	243	16
10^{-25}	299	18

6 结束语

文中构建了一类基函数神经网络模型,定义了其优化目标函数,且依据梯度下降法原理,推导出该类神经网络学习算法.在此基础上,构造了该类神经

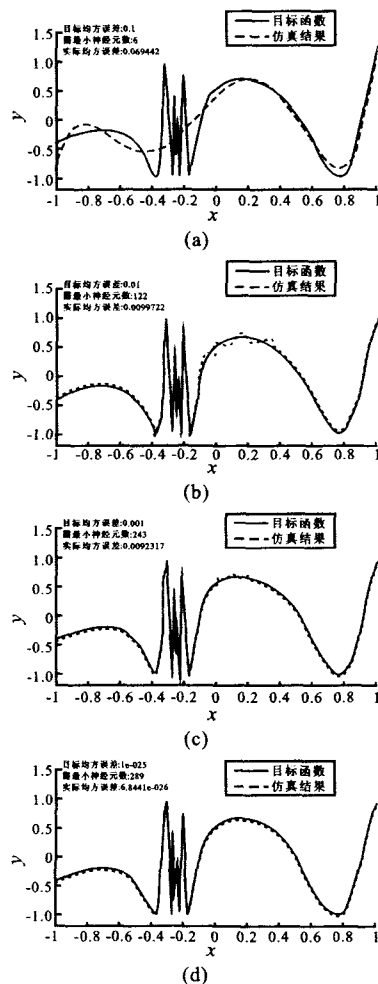


图2 广义多项式神经网络及其结构自适应确定结果

网络最小隐层结构的自适应、快速确定算法.仿真实验表明,文中算法是可行和有效的.

参考文献:

- [1] 汪烈军.一种改进的结构自适应自组织神经网络算法[J].微电子学与计算机,2007,24(1):106-109.
- [2] Zhang Y, Zhong T, Li W, et al. Growing algorithm of laguerre orthogonal basis neural network with weights directly determined[C]//Proc. ICIC08. Berlin: Springer Press, 2008:60-67.
- [3] 乔俊飞,张颖.一种多层前馈神经网络的快速修剪算法[J].智能系统学报,2008,3(4):174-176.
- [4] 崔荣一,洪炳熔.关于三层前馈神经网络隐层构建问题的研究[J].计算机研究与发展,2004,41(4):524-530.
- [5] 邹阿金,成继勋,沈建中.基于Chebyshev神经网络的衍生算法[J].湘潭矿业学院学报,2001,16(1):76-78.
- [6] Fahlman S E, Lebiere C. The cascade-correlation learning architecture[C]//Advances in Neural Information Processing Systems. USA, San Mateo, 1990.

- [7] 李倩,王永县,朱友芹. 人工神经网络混合剪枝算法[J]. 清华大学学报, 2005, 45 (6): 831 - 834.
- [8] Islam M, Murase K. A new algorithm to design compact two-hidden-layer artificial neural networks[J]. Neural Networks, 2001(14): 1265 - 1278.
- [9] Huang G B, Saratchandran P, Sundararajan N. A generalized growing and pruning RBF (CGAP-RBF) neural network for function approximation[J]. IEEE Trans. on Neural Networks, 2005(16): 57 - 67.
- [10] Ahmed S, Abdullah - Al - Mamun K, Islam M. A novel algorithm for designing three layered artificial neural networks[J]. International Journal of Soft Computing, 2007, 2(3): 450 - 458.
- [11] Narasimhaa P L, Delashmitb W H, Manrya M T, et al. An integrated growing - pruning method for feedforward network training[J]. Neuro computing, 2008(71): 2831 - 2847.

作者简介:

肖秀春 男, (1976 -), 博士研究生, 讲师. 研究方向为神经网络、图像处理、信息安全等.

姜孝华 男, (1963 -), 博士, 教授, 博士生导师. 研究方向为神经网络、信号检测与处理技术.

张雨浓 男, (1973 -), 博士, 教授, 博士生导师. 研究方向为神经网络、机器人、高斯过程.

(上接第 56 页)

- [4] IEEE Standard. Binary floating - point arithmetic[S]. ANSI/IEEE Standard 754, 1985.
- [5] Schmookler M S, Nowka K J. Leading zero anticipation and detection - a comparison of methods[C]// Proc. IEEE 15th symp. USA: Austin Computer Arithmetic, 2001: 7 - 12.
- [6] 梅小露. 浮点乘加部件中三操作数前导 1 预测算法的设计[J]. 微电子学与计算机, 2005, 22(12): 16 - 20.
- [7] Javier B, Tomas L. Leading-one prediction with concurrent position correction[J]. IEEE Trans. Computers, 1999, 48(10): 1083 - 1097.
- [8] Yu R K, Zyner G B. 167 MHz radix 4 floating point multiplier[C]// Proc. IEEE 12th Symp. Computer Arithmetic. UK, Bath, 1995: 149 - 154.

作者简介:

蔡 敏 男, (1955 -), 博士, 教授, 博士生导师. 研究方向为集成电路设计.

闵言灿 男, 硕士. 研究方向为集成电路设计.