用最陡下降法修正RBF隐层参数

李　波[1,2]，王成友[2]，蔡宣平[2]，唐朝京[2]，张尔扬[2]
(1. 空军工程大学 电讯工程学院，陕西 西安 710077;2. 国防科技大学 电子科学与工程学院，湖南 长沙 410073)

摘　要:在用传统聚类方法得到初步的RBF隐节点参数之后,提出再用最陡下降法进行误差反传学习,进一步校正隐节点参数。仿真实验证明该方法可以使RBF网络的函数逼近能力明显增强。

关键词:RBF神经网络;最陡下降法;全局逼近
中图分类号:TP183　**文献标识码**:A　**文章编号**:1009-3516(2004)04-0067-03

RBF(Radial Basis Function Neural Network)神经网络模拟了在人的大脑皮层区域中,局部调节及感受域(Receptive Field)交叠的人脑反应特点,是一种性能良好的前向神经网络模型,它具有全局逼近的性质,且不存在局部最小问题[1,4,5]。RBF是具有单隐层的3层前馈网络,其输入层到隐含层的变换是非线性的,而从隐含层到输出层的变换是线性的。RBF网络隐含层的转移函数通常取为高斯函数、细薄板样条函数等[2],其隐层中心参数和宽度的选取一直是研究的关键,研究者们提出了不少选取方法,如随机选取法、k-均值法、Konhonen自组织映射法、最近邻聚类选取法等。研究表明,最近邻聚类选取法具有学习时间短,计算量小和网络性能优良的优点,是一种实用的前馈网络的快速学习算法。但这里隐层中心参数和密度的选取完全依据输入变量来确定,而隐层到输出层间的连接权值则由最小二乘法来学习确定,这种分开的学习方法虽然学习速率快,可往往隐层中心参数和宽度选取不一定准确,由此导致最终的结果不准确。

本文提出先由通常的聚类方法得到隐层中心参数和宽度的初值,再通过最陡下降法来进行误差反传学习,调整隐层到输出层的连接权值以及隐层中心参数和密度,以此来提高RBF网络的函数逼近能力。实验结果表明,通过最陡下降法调整后,函数的逼近能力得到明显的提高。

1　RBF神经网络

RBF神经网络的典型结构如图1所示,不失一般性,假设输入层和输出层只有1个节点,这种结构很容易扩展到多输入、多输出节点的情形,输入层到隐层为权值1的固定连接。隐层由1组径向基函数构成,与每个隐层节点相关的中心参数为c_i,宽度参数为σ_i,隐层至输出层连接权值为ω_i,此处研究的径向基函数取高斯函数。

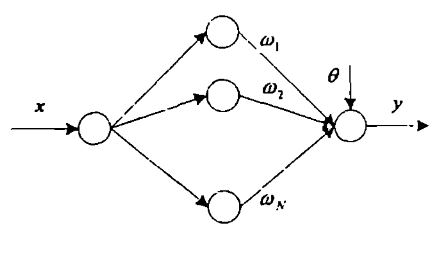


图1　RBF神经网络结构

$$y=f(x)=\sum_{i=1}^{n}\omega_i\cdot node_i-\theta=\sum_{i=0}^{n}\omega_i\cdot node_i \tag{1}$$

$$node_i=\exp\left|-\frac{\|x-c_i\|^2}{2\sigma_i^2}\right| \tag{2}$$

$node_i$和c_i为第i个隐节点的输出值和中心值，θ为输出单元偏移量，ω_i为第i个隐节点到输出节点的连接权值，$\|\cdot\|$为欧几里得泛数，$\omega_0=\theta,node_0=-1$。构造和训练1个RBF神经网络就是要使它通过学习,确定出每个隐层神经元基函数的中心c_i和宽度σ_i,以及隐层到输出层的权值ω_i。

收稿日期:2003-12-26
作者简介:李　波(1974-),男,山东青岛人,博士生,主要从事语音信号处理及模式识别研究.

2　网络参数的学习

2.1　k－均值聚类算法初步确定 c_i 和 σ_i

k－均值聚类算法是一种模式聚类法,其原理是首先初始化中心 c_i,然后将输入样本按最邻近规则分组,即将 $x_j(j=1,2,\cdots,H)$ 分配给中心 $c_i(i=1,2,\cdots,n)$ 的输入样本聚类集合 $\varphi_i(i=1,2,\cdots,n)$ 。要求满足:

$$d_i^j=\min_i \| x_j-c_i \| \quad i=1,2,\cdots,n;\quad j=1,2,\cdots,H \tag{3}$$

然后,根据样本 φ_i的平均值调整中心 c_i。重复以上步骤,直到中心的分布不再变化。中心确定后,再相应地计算出宽度 σ_i。

2.2　用最陡下降法对 ω_i进行学习

对于由 H 个输入/输出样本对 $x_i/d_i(i=1,2,\cdots,H)$ 组成的一组训练数据,定义目标函数:

$$E=\frac{1}{2}\sum_{i=1}^{H}(d_i-y_i)^2$$

$$\frac{\partial E}{\partial \omega_i}=-(d_i-y_i)\frac{\partial y_i}{\partial \omega_i}=-(d_i-y_i)node_i \tag{4}$$

$$\omega_i(t+1)=\omega_i(t)-\eta\frac{\partial E}{\partial \omega_i} \tag{5}$$

η 为学习速率。

2.3　用最陡下降法对 c_i 和 σ_i 进行再学习

$$\frac{\partial E}{\partial c_i}=-(d_i-y_i)\frac{\partial y_i}{\partial node_i}\frac{\partial node_i}{\partial c_i}=-(d_i-y_i)\omega_i\exp\left(-\frac{(x_i-c_i)^2}{2\sigma_i^2}\right)\frac{(x_i-c_i)}{\sigma_i^2} \tag{6}$$

$$c_i(t+1)=c_i(t)-\eta\frac{\partial E}{\partial \sigma_i} \tag{7}$$

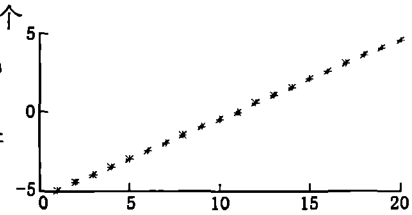
$$\frac{\partial E}{\partial \sigma_i}=-(d_i-y_i)\frac{\partial y_i}{\partial node_i}\frac{\partial node_i}{\partial \sigma_i}=-(d_i-y_i)\omega_i\exp\left(-\frac{(x_i-c_i)^2}{2\sigma_i^2}\right)\frac{(x_i-c_i)^2}{\sigma_i^3} \tag{8}$$

$$\sigma_i(t+1)=\sigma_i(t)-\eta\frac{\partial E}{\partial \sigma_i} \tag{9}$$

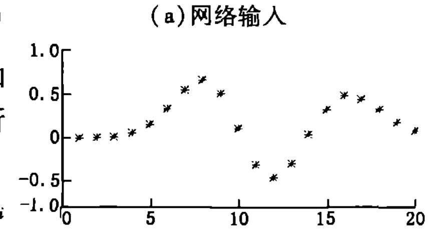
3　仿真研究

设网络的输入为 $x=-5\cdot 0.5\cdot 4.5$,目标为单输入/单输出、隐层有 3 个节点的高斯 RBF 网络确定。3 个隐节点的中心分别为 $c_1=-1$、$c_2=0$、$c_3=2.5$;标准化参数 $\sigma_i^2=1(i=1,2,3)$;隐层到输出层的权分别为 $\omega_1=1.1$、$\omega_2=-1$、$\omega_3=0.5$。输入与输出关系由式(10)表示,关系如图 2 所示。

$$d_i=1.1\exp\left(-\frac{(x_i+1)^2}{2}\right)-\exp\left(-\frac{x_i^{\,2}}{2}\right)+0.5\exp\left(-\frac{(x_i-2.5)^2}{2}\right) \tag{10}$$



(a)网络输入



(b)网络输出

图 2　目标输入与输出关系图

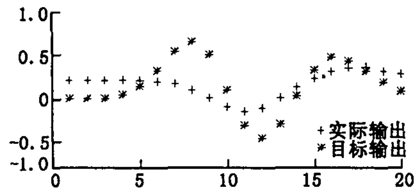
实验 1:采用原始的方法,即用 k－均值法对输入进行聚类,求 c_i 和 σ_i;保持 c_i和 σ_i 不变,用最陡下降法对 ω_i进行学习,得到结果如图 3 所示。

实验 2:采用本文的提出的方法,用 k－均值对输入进行聚类,求 c_i 和 σ_i;保持 σ_i 不变,用最陡下降法训练 c_i 和 ω_i,得结果如图 4 所示。

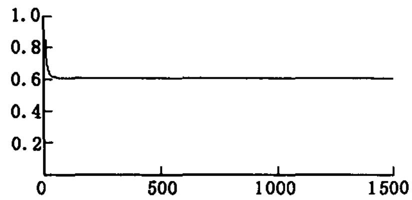
实验 3:用 k－均值对输入进行聚类,求 c_i 和 σ_i;再用最陡下降法训练 c_i、σ_i 和 ω_i,得结果如图 5 所示。

从图3、图4 和图5 的实验结果可以看出,用误差反传的方法对 c_i、σ_i 和 ω_i 进行再学习效果是明显的,当只对 ω_i学习,误差达 0.6 左右;而对 c_i和 ω_i进行学习就使误差达到 0.1 以下;对 c_i、σ_i 和 ω_i 一起进行学习,使

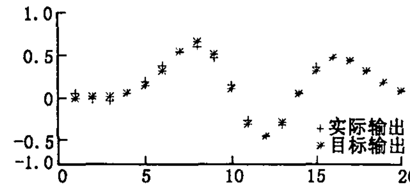
误差更小。



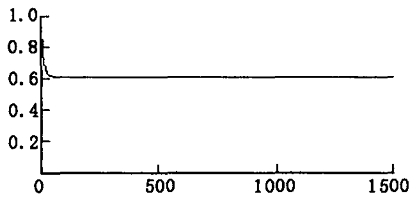
(a)网络输入



(a)网络输入

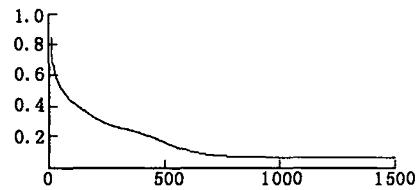


(a)网络输入



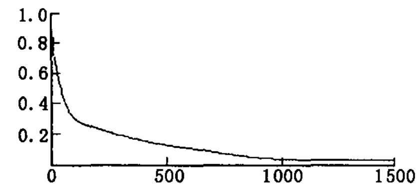
(b)网络输出

图3 对 ω_i 进行学习结果



(b)网络输出

图4 对 c_i、ω_i 进行学习结果



(b)网络输出

图5 对 c_i、σ_i 和 ω_i 进行学习结果

RBF 网络的输出误差函数也为 c_i、σ_i 和 ω_i 的函数,受 BP 网络误差反传学习算法的启发,提出在进行常规的 RBF 训练之后,用误差反传算法再进行系数的进一步调整,仿真实验结果表明本算法有效地提高了函数逼近精度。

参考文献:

[1] Binchini M, Frasconi P, Gori M. Learning Without Local Minima in Radial Basis Function Networks[J]. IEEE Transaction on Neural Networks, 1995, 6(3):749-755.

[2] 李鑫滨,杨景明,丁喜峰. 基于递推 k-均值聚类算法的 RBF 神经网络及其在系统辨识中的应用[J]. 燕山大学学报, 1999, 23(4): 363-366.

[3] 朱明星,张德龙. RBF 网络基函数中心选取算法的研究[J]. 安徽大学学报(自然科学版), 2003, 24(1): 72-78.

[4] 王旭东,邵惠鹤. RBF 神经网络理论及其在控制中的应用[J]. 信息与控制, 1997, 26(4): 272-284.

[5] 李彦民,白本督,李 映. 前向神经网络拓扑结构和权值的进化[J]. 空军工程大学学报(自然科学版),2001,2(4):76-76.

(编辑:门向生)

Modifying the Hidden Layer Parameters by Using the Method of Steepest Descent

LI Bo[1,2], WANG Cheng-you[2], CAI Xuan-ping[2], TANG Chao-jing[2], ZHANG Er-yang[2]

(1. The Telecommunication Engineering Institute, Air Force Engineering University, Xi'an, Shaanxi 710077, China; 2. College of Electronic Science and Engineering, National University of Defense Technology, Changsha, Hunan 410073, China)

Abstract: At first, the hidden node parameters of RBF are obtained by the classical clustering method; then those parameters are studied again by the method of steepest descent. The simulation shows that the method can reinforce the RBF approximation quality obviously.

Key words: RBF neural networks; method of steepest descent; universal approximation