

文章编号:1005-0523(2009)03-0037-05

径向基神经网络辨识非线性系统的改进算法

孟莎莎, 孙宝华, 杨 辉, 袁路生

(华东交通大学 电气与电子工程学院, 江西 南昌 330013)

摘要:利用神经网络的非线性映射特性,将神经网络应用于非线性系统辨识。利用径向基神经网络来辨识非线性系统,并对两种不同 RBF 神经网络辨识算法进行比较。仿真结果表明,改进的算法具有学习速度快,辨识精度高的特点。

关键词:径向基神经网络;非线性系统;系统辨识

中图分类号:TP13

文献标识码:A

人工神经网络特有的非线性适应信息处理能力,克服了传统人工智能方法对直觉,如模式、语音识别、非结构化信息处理方面的缺陷,使之在神经专家系统、模式识别、智能控制、组合优化、预测等领域得到成功应用。近年来,人工神经网络正向模拟人类认知的道路上更加深入发展,与模糊系统、遗传算法、进化机制等结合,形成计算智能,成为人工智能的一个重要方向,将在实际应用中得到发展^[1]。

本文就是利用径向基(Radial Basis Function, RBF)神经网络来辨识非线性系统,并对两种不同 RBF 神经网络辨识算法进行比较。仿真结果显示,改进的算法具有学习速度快,辨识精度高的特点。

1 辨识算法

1.1 RBF 神经网络

RBF 网络的结构与多层前向网络类似,但它是具有单隐含层的一种两层前向网络。输入层由信号源节点组成,隐含层的单元数视所描述问题的需要而定,输出层对输入的作用做出响应。从输入空间到隐含层空间的变换是非线性的,而从隐含层空间到输出层空间的变化是线性的。隐单元的变化函数是 RBF,它是一种局部分布的对中心点径向对称衰减的非负非线性函数。

RBF 网络由两层组成,其结构如图 1 所示。输入层节点只是传递输入信号到隐含层,隐含层节点由像高斯函数那样的辐射状作用函数构成,而输出层节点通常是简单的线性函数。

隐含层节点中作用函数对输入信号将在局部产生响应,也就是说,当输入信号靠近该函数的中央范围时,隐含层节点将产生较大的输出。由此可看出这种网络具有局部逼近能力,故径向基函数网络也称为局部感知场网络。

隐含层节点的作用函数有多种形式,但最常用的是高斯函数,因为高斯函数具有以下优点:

- (1) 径向对称,形式简单,易于描述问题。
- (2) 光滑性好,任意阶导数存在。
- (3) 解析性好,因而便于进行理论分析。

隐含层第 i 个节点的输出为

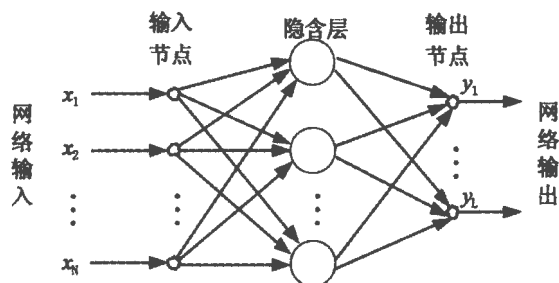


图 1 RBF 网络结构

收稿日期:2009-03-18

基金项目:国家自然科学基金项目(60864004);江西省教育厅项目(GJJ09222);华东交通大学研究生创新基金项目(YC08004)

作者简介:孟莎莎(1984-),女,山西大同人,硕士研究生,研究方向为复杂系统的非线性控制与预测控制。

$$h_i(x) = \exp\left(-\frac{(x - c_i)^T(x - c_i)}{2\sigma_i^2}\right), (i = 1, 2, \dots, q) \quad (1)$$

网络输出为

$$y(x) = \sum_{i=1}^M w_i h_i(x) \quad (2)$$

其中, $x = (x_1, x_2, \dots, x_M)^T$ 是输入向量, $c_i = (c_{i1}, c_{i2}, \dots, c_{iM})^T$ 是第 i 个隐节点高斯函数中心向量, 此向量是一个与输入维数相同的列向量, M 是隐含层节点数, σ_i 是第 i 个隐节点高斯函数形状参数。

从整体看, 网络由输入到输出是非线性的, 从局部看, 网络输出对隐含层却是线性的, 可以通过调整权系数来改变网络的输出, 从而加快了学习速度, 避免局部极小的问题。

1.2 常规算法

RBF 网络的学习过程分两个阶段。第一个阶段是无教师学习, 是根据所有的输入样本决定隐含层各节点的高斯函数的中心向量 c_i 和标准化常数 σ_i 。第二阶段是有教师学习。在决定好隐含层的参数后, 根据样本, 利用最小二乘原则, 求出隐含层和输出层的权值 w_{ki} 。下面具体介绍这两个阶段。

1) 无教师学习阶段

无教师学习也称为非监督学习, 是对所有样本的输入进行聚类, 求得各隐含层节点 RBF 的中心向量 c_i 。这里介绍 k 均值聚类算法调整中心向量^[1,2], 此算法将训练样本集中的输入向量分为若干族, 在每个数据族内找一个径向基函数中心向量, 使得该族内各样本向量距该族中心的距离最小。算法步骤如下:

(1) 给定各隐节点的初始中心向量 $c_i(0)$ 和判断停止计算的 ϵ ;

(2) 计算距离(欧氏距离)并求出最小距离的节点

$$\begin{cases} d_i(k) = \|x(k) - c_i(k-1)\|, 1 \leq i \leq q \\ d_{\min}(k) = \min d_i(k) = d_r(k) \end{cases} \quad (3)$$

式中, k 为样本序号, r 为中心向量 $c_i(k-1)$ 与输入样本 $x(k)$ 距离最近的隐节点序号;

(3) 调整中心

$$\begin{cases} c_i(k) = c_i(k-1), 1 \leq i \leq q, i \neq r \\ c_r(k) = c_r(k-1) + \beta(k)(x(k) - c_r(k-1)) \end{cases} \quad (4)$$

式中, $\beta(k)$ 是学习速率。 $\beta(k) = \beta(k-1)/(1 + \text{int}(k/q))^{1/2}$, $\text{int}(\cdot)$ 表示对 $(\cdot)$ 进行取整运算。可见, 每经过 q 个样本之后, 调小一次学习速率, 逐渐减至为零;

(4) 判定聚类质量

对于全部样本 $k(k = 1, 2, \dots, N)$ 反复进行以上②, ③步, 直至满足下述条件, 则聚类结束。

$$J_e = \sum_{i=1}^q \|x(k) - c_i(k)\|^2 \leq \epsilon \quad (5)$$

2) 有教师学习阶段

有教师学习也称为有监督学习。当 c_i 确定以后, 训练由隐含层至输出层之间的权值, 由上可知, 它是一个线性方程组, 则求权值就成为线性优化问题。因此, 问题有唯一确定的解, 不存在 BP 网络中所遇到的局部极小值问题, 肯定能获得全局最小点。类似于线性网络, RBF 网络的隐含层至输出层之间的连接权值 w_{ki} 学习算法为

$$w_{ki}(k+1) = w_{ki}(k) + \eta(t_k - y_k)u_i(x)/u^T u \quad (k = 1, 2, \dots, L; i = 1, 2, \dots, q) \quad (6)$$

式中, $u = [u_1(x), u_2(x), \dots, u_q(x)]^T$, $u_i(x)$ 为高斯函数; $\eta(0 < \eta < 1)$ 为学习速率; t_k 和 y_k 分别表示第 k 个输出分量的期望值和实际值。

1.3 改进算法

设有多输入单输出非线性对象为

$$y(k+1) = f[y(k), y(k-1), \dots, y(k-n), u(k), u(k-1), \dots, u(k-m)] \quad (7)$$

所要完成的任务是:建立 RBF 神经网络,通过学习训练,逼近函数 $f(\cdot)$,完成动态建模,使实际输出 $y(k+1)$ 与网络输出 $\hat{y}(k+1)$ 一致。

取网络输入向量为

$$\begin{aligned} x(k) &= [x_1(k), x_2(k), \dots, x_N(k)]^T \\ &= [y(k), y(k-1), \dots, y(k-n), u(k), u(k-1), \dots, u(k-m)]^T \end{aligned} \quad (8)$$

其中 $N = n + m + 2$ 。隐含层第 i 个节点的输出同式(1)。网络输出为

$$\hat{y}(k+1) = \sum_{i=1}^M w_i(k) h_i(x(k)) \quad (9)$$

其中, M 为隐含层神经元个数,应由具体情况而定,一般有 $M > N$ 。定义目标函数为

$$J = \frac{1}{2} (y(k+1) - \hat{y}(k+1))^2 \quad (10)$$

学习过程按使误差函数 J 减小最快的方向调整加权系数直到获得满意的加权系数集为止。根据梯度法,可得隐含层至输出层的加权系数修整公式为

$$w_i(k+1) = w_i(k) - \alpha_i \frac{\partial J}{\partial w_i(k)} \quad (11)$$

根据寻求中间变量求偏导的方法,得

$$\frac{\partial J}{\partial w_i(k)} = \frac{\partial J}{\partial y(k+1)} \frac{\partial y(k+1)}{\partial w_i(k)} = -(y(k+1) - \hat{y}(k+1)) h_i(k) \quad (12)$$

联合式(11)和式(12)得最后的调整公式为

$$w_i(k+1) = w_i(k) + \alpha_i (y(k+1) - \hat{y}(k+1)) h_i(k) \quad (13)$$

由前小节的分析,可知采用高斯函数作为基函数的 RBF 神经网络是局部逼近网络,实践中发现,高斯函数的形状参数和中心向量对网络输出影响很大,在它们固定不变的前提下,只有当输入落在中心向量附近时,才对输出有影响。

前面介绍了 RBF 网络的两个学习阶段,第一阶段是无教师学习,是根据所有的输入样本决定隐含层各节点的高斯函数的中心向量 c_i 和形状参数 σ_i ;第二阶段是有教师学习,在决定好隐含层的参数后,根据样本,利用最小二乘原则,求出隐含层至输出层的权值 w_i 。这样计算量就较大,而且中心向量和形状参数确定以后,单靠调整连接权系数来逼近非线性,有时会出现收敛时间过长,逼近精度差的现象。为了解决这个问题,也有采用增加神经元数目的办法来解决,这样又增加了计算量。现在增加调整基函数中心向量和形状参数的做法,它既不增加隐含层神经元数目,又可以提高逼近速度和精度^[3]。根据梯度法,中心向量和形状参数调整公式为^[4]

$$\begin{aligned} c_i(k+1) &= c_i(k) + \beta_i \frac{\partial J}{\partial c_i(k)} \\ &= c_i(k) - \beta_i \frac{\partial J}{\partial y(k+1)} \frac{\partial y(k+1)}{\partial h_i(k)} \frac{\partial h_i(k+1)}{\partial c_i(k)} \frac{\partial h_i(k)}{\partial c_i(k)} \\ &= c_i(k) - \beta_i (y(k+1) - \hat{y}(k+1)) w_i(k) h_i(k) (x - c_i(k)) / \sigma_i^2 \end{aligned} \quad (14)$$

$$\begin{aligned} \sigma_i(k+1) &= \sigma_i(k) + \gamma_i \frac{\partial J}{\partial \sigma_i(k)} \\ &= \sigma_i(k) - \gamma_i \frac{\partial J}{\partial y(k+1)} \frac{\partial y(k+1)}{\partial h_i(k)} \frac{\partial h_i(k)}{\partial \sigma_i(k)} \\ &= \sigma_i(k) - \gamma_i (y(k+1) - \hat{y}(k+1)) w_i(k) h_i(k) \| (x - c_i(k)) \| / \sigma_i^3 \end{aligned} \quad (15)$$

式中, α_i 、 β_i 和 γ_i 为学习速率,一般 $\in (0, 1]$ 用于控制权值调整速度。

2 仿真比较

现用式(16)所表示的 Hermite 函数^[5,6]来进行仿真比较:

$$y = 1.1 \times (1 - x + 2x^2) \exp\left(-\frac{1}{2}x^2\right) \quad (16)$$

仿真效果如图2和图3所示。

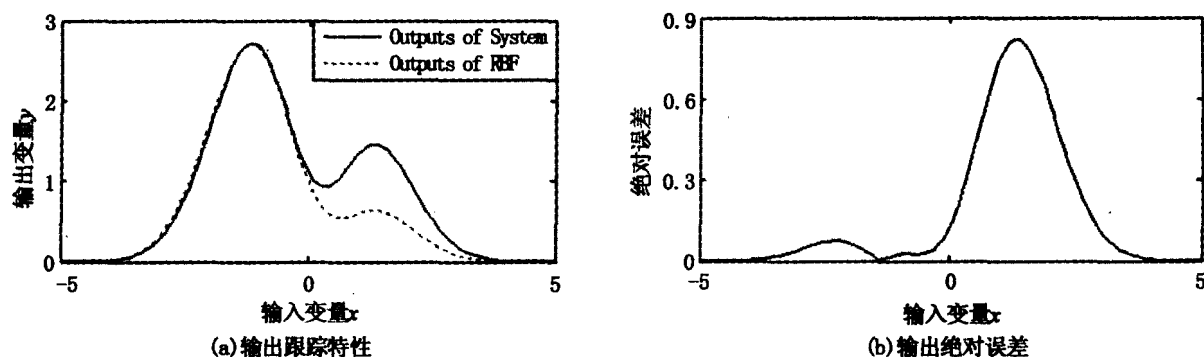


图2 采用算法1的仿真结果

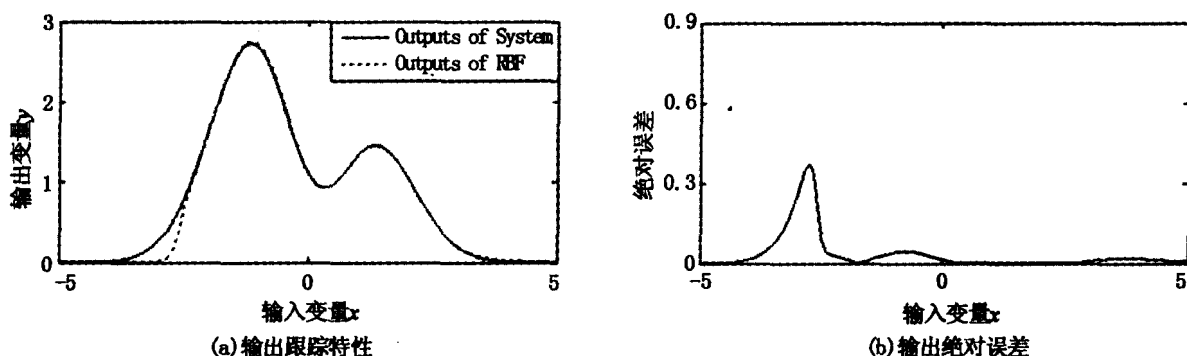


图3 采用算法2的仿真结果

比较可见,采用算法2的辨识效果优于算法1的辨识效果。算法2开始出现的误差是由于开始还处于边学习,边辨识过程。同时不难看出,采用算法2,RBF网络具有学习速率快,辨识精度高的优点。

3 结论

本文提出了一种将学习和辨识同时进行的在线辨识技术,并且在调整输出权值的基础上,增加调整基函数的形状参数和中心向量。与传统的RBF算法相比,改进的RBF网络算法学习速度快,适于在线实时控制。

参考文献:

- [1] 李国勇.智能控制及其MATLAB实现[M].北京:电子工业出版社,2005.
- [2] 杨辉,谭明皓,柴天佑.基于神经网络的多元稀土萃取组分含量的软测量[J].中国稀土学报,2003,21(4):425-430.
- [3] 徐湘元,毛宗源.基于径向基函数神经网络的内膜控制[J].电路与系统学报,1999,4(2):86-91.
- [4] 徐湘元.自适应控制理论与应用[M].北京:电子工业出版社,2007.
- [5] 王学雷,邵惠鹤,李亚芬.一种径向基函数神经网络在线训练算法及其在非线控制中的应用[J].信息与控制,2001,30(3):249-253.
- [6] Narendra K, Gallman P. An iterative method for the identification of nonlinear systems using a Hammerstein model[J]. IEEE Trans on Automatic Control, 1966, 11(3):546-550.

Improved Algorithm of Nonlinear System Identification Using RBF Neural Network

MENG Sha-sha, SUN Bao-hua, Yang Hui, ZHONG Lu-sheng

(School of Electrical and Electronic Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: Using the nonlinear-mapping features of neural network, a method for nonlinear system identification using neural network is presented in this paper. It uses RBF neural network to identify and compares two different algorithms of RBF neural network. The simulation results indicate that the new algorithm of nonlinear system identification is fast to learn and precise to identify.

Key words: RBF neural network; nonlinear system; system identification

(责任编辑:王建华)

(上接第25页)

On Design Method of Temporary Pavement and Bridge System by the Cover-excavation Method

ZHAN Chao, YANG Long-cai, HE Tai-hong

(Key Laboratory of Road and Traffic Engineering in Education Ministry, Tongji University, Shanghai, 201804, China)

Abstract: In order to minimize the influence of subway station construction on the existing ground traffic, the cover-excavation method is widely used. In this case, a keypoint is to build a stable and reliable temporary road and bridge system. There is no specific design specification in this field. Combining with a project of metro station constructed by cover-excavation method, the design calculation of temporary pavement slab, the parametric configuration of military beam system, the interaction of military beam and enclosure structure are studied. Many useful conclusions are obtained to provide general principles of design calculation for temporary road and bridge system by cover-excavation method.

Key words: cover-excavation method; temporary road and bridge system; pavement slab; military beam; numerical simulation

(责任编辑:刘棉玲)