

改进的 RBF 神经网络在非线性系统中的应用

储岳中

(安徽工业大学 计算机学院, 安徽 马鞍山 243002)

摘要:在 RBF 神经网络的各种学习算法中,最近邻聚类算法学习时间短、计算量小,不需要事先确定隐单元的个数,完成聚类所得到的网络是最优的,并且可以在线学习,是一种自适应聚类学习算法,非常适合非线性实时系统的应用。但常规最近邻聚类算法在实时性要求较高的系统预测中学习时间相对较长。针对这一问题,提出了系统离线学习时采用减聚类算法,在线学习时采用改进的最近邻聚类算法,并变步长修正聚类半径和限制学习样本数。在函数拟合实验中,这种改进算法明显缩短了 RBF 神经网络的学习时间,在钢包精炼炉电极系统的在线辨识中的成功应用进一步表明对最近邻聚类算法的改进是有效的。

关键词:RBF 神经网络;减聚类算法;最近邻聚类算法;系统辨识;钢包精炼炉

中图分类号:TP183

文献标识码:A

文章编号:1673-629X(2008)03-0196-04

Application of an Improved RBFNN in Nonlinear System

CHU Yue-zhong

(School of Computer, Anhui University of Technology, Ma'an Shan 243002, China)

Abstract: The nearest neighbor-clustering algorithm has a short training time, less work to calculate and the number of hidden units is not to be determined in advance in the various RBFNN learning algorithms, the network is optimization after clustering and can be trained on-line, it is an adaptive clustering algorithm for nonlinear real-time system. But it needs a longer training time in a real-time system which require high precision. Aimed at this problem, an improved algorithm was presented by using subtractive clustering algorithm in off-line learning and nearest neighbor-clustering algorithm in on-line learning, the clustering radius was modified by using variable step, the example number was limited in on-line identification. This algorithm cuts training time of RBFNN a lot in function fitting. The application in on-line identification of ladle furnace electrode system indicates the efficiency of the improved algorithm.

Key words: RBFNN; subtractive clustering algorithm; nearest neighbor-clustering algorithm; system identification; ladle furnace

0 引言

RBF(径向基函数)神经网络是一种两层的前馈网络,有唯一的隐层。隐节点的输出是径向基函数,这种基函数对输入激励产生一个局部化的响应,即仅当输入落在输入空间中的一个很小的指定区域中时,隐单元才做出有意义的响应,网络的输出是所有隐单元 c_i 输出的线性变换。RBF 神经网络的网络结构如图 1 所示。RBF 网络的输出采用加权平均法推导时^[1],可得式(1)形式:

$$f(X_k) = \frac{\sum_{i=1}^{c_n} W_i R_i(X_k)}{\sum_{i=1}^{c_n} R_i(X_k)} \quad (1)$$

其中, $R_i(X_k)$ 为网络第 i 个隐结点的输出,当基函数取高斯函数时可用式(2)计算:

$$R_i(X_k) = \exp\left(-\frac{\|X_k - C_i\|^2}{r^2}\right) \quad (2)$$

式(1)和(2)中, $X_k = (x_1, x_2, \dots, x_n)$ 为第 k 个输入向量, C_i 为聚类中心向量, W_i 为第 i 个隐结点到输出结点的权值, r 为高斯函数的宽度。

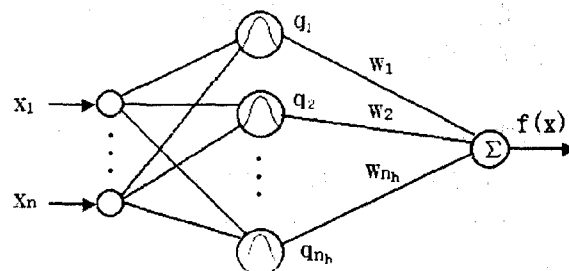


图1 单输出的 RBF 神经网络结构

收稿日期:2007-06-15

基金项目:安徽省高校青年教师科研资助计划项目(2006jq089)

作者简介:储岳中(1971-),男,安徽岳西人,硕士,讲师,主要从事人工智能及计算机仿真方面的研究。

1 算法分析

构造和训练一个 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 , 宽度 r 以及隐层到输出层的权值 W_i , 这些参数的过程, 从而可以完成所需的输入到输出的映射。由于 RBF 神经网络中隐层结点与输出结点所起的作用不同, 其训练方法大致可分为两类^[2]:

①径向基函数的中心、宽度与隐层到输出层的连接权重分别进行训练。

②径向基函数的中心、宽度与隐层到输出层的连接权重同时进行训练。

文中采用的是第一类训练方法, 基函数的中心和宽度在不同阶段分别用减聚类算法和最近邻聚类算法确定, 中心和宽度确定后, 网络输出对输出层权参数是线性的, 可以采用类内不同数据点的输出分量平均求得。

RBF 神经网络现有算法一般可分为离线算法和在线算法。离线方式的训练算法是一种定结构的 RBF 网络, 即从训练样本空间的分布模式中提取中心, 中心选定之后, 网络结构也就确定下来了, 当数据模式改变时, 网络无法作相应改变, 因此离线方式不适用于时变系统。所谓在线方式, 是将中心选取、参数校正集于一体, 在输入样本数据的同时, 选择中心并校正网络参数, 动态地构建网络的拓扑结构, 这类算法兼有实时性和快速性的优点。

针对非线性实时系统, 为了实现网络的在线辨识并获得良好的网络辨识性能, 文中将两种方法结合起来: 首先采集一批数据用于批量处理模式离线学习, 得到一组比较好的网络权值, 构建 RBF 神经网络的初始结构; 再采用在线辨识, 即采集一组新的样本, 网络参数就修正一次。由于有了较好的网络初值, 网络的性能会大大提高。具体的做法就是先利用减聚类算法为网络确定初值, 再用动态最邻近聚类算法实现在线辨识。

2 改进的最近邻聚类算法

2.1 网络的离线学习

Chiu 提出的减聚类方法^[2]是一种相对简单而又有效的聚类算法。在该算法中, 数据集自身作为聚类中心的候选集。用此方法, 计算量与数据点的数目表现为简单的线性关系, 且与所考虑问题的维数无关。解决了一般聚类算法计算量随问题的维数增加而呈指数增长的问题, 也即“维数灾难”。考虑 n 维空间的 p 个数据点 (x_1, x_2, \dots, x_p) , 假定数据已归一化到一个单位超立方体中, 首先给出数据点 x_i 处的密度指标定义^[3]:

$$D_i = \sum_{j=1}^p \exp \left[- \frac{\|x_i - x_j\|^2}{(\delta_1/2)^2} \right] \quad (i = 1, 2, \dots, p) \quad (3)$$

在计算每个数据点密度指标后, 选择具有最高密度指标的点作为第一个聚类中心, 其密度指标记为 D_{c_1} 。采用(4)式更新每个数据点的密度指标:

$$D_i = D_i - D_{c_1} \sum_{j=1}^p \exp \left[- \frac{\|x_i - x_{c_m}\|^2}{(\delta_2/2)^2} \right] \quad (i = 1, 2, \dots, p) \quad (4)$$

取常数 $\delta_2 > \delta_1$ (根据经验取 $\delta_2 = 1.5\delta_1$ ^[4]), 这样在其邻域内, 密集度指标显著减小, 可以避免出现相距很近的聚类中心。显然, 靠近第一个聚类中心 x_{c_m} 的数据点的密度指标将显著减小, 这样使得这些点不太可能成为下一个聚类中心。更新每个点的密度指标后, 选定下一个聚类中心, 再次修正密度指标, 此过程不断重复, 直到满足当前最高密度指标 D_{\max} 同初始最高密度指标相比非常小, 即:

$$\frac{D_{\max}}{D_{c_1}} < \lambda \quad (5)$$

由此结束聚类, 所得到的聚类中心个数即为基函数中心个数。具体的学习算法如下:

Step1 按式(3)计算每个数据点的密度指标, 将 D_i 保存到集合 $A(c)$ 中, 供下次迭代用, 选择具有最高密度指标 D_{c_1} 的数据点 x_{c_1} 为第一个聚类中心, 记类数 $m = 1$ 。

Step2 确定第 m 个聚类中心。利用式(4)修正每个数据点的密度指标, 寻找最高密度指标 D_{\max} 。

Step3 将 D_{\max} 代入式(5), 如果成立, 则不接受该点为聚类中心, 转至 Step4; 否则, 接受具有最高密度指标的数据点 x_{c_m} 为第 m 个聚类中心, 并将 D_i 保存到集合 $A(c)$ 中, 且 $D_{c_m} = D_i$, $m = m + 1$, 返回 Step2。

Step4 聚类结束, 显然共确定 $m - 1$ 个聚类中心 $x_{c_1}, x_{c_2}, \dots, x_{c_{m-1}}$, 记 $c = m - 1$ 。

确定了聚类中心后, 对每个类内所有数据点取平均距离即可确定基函数宽度 δ_i , 对每个类内所有数据点向量中的输出分量平均即可确定输出层的权值 W_i ^[1]。

2.2 网络的在线学习

RBF 神经网络算法的核心是基函数个数、中心的选取与宽度的确定。动态最近邻聚类算法学习时间短、计算量小, 不需要事先确定隐单元的个数, 完成聚类所得到的网络是最优的, 并且可以在线学习, 是一种自适应聚类学习算法, 非常适合非线性实时系统的应用。

系统离线学习后已确定一个基本的 RBF 神经网络结构, 在此基础上的在线学习算法步骤如下:

Step1 计算所有离线学习时数据点间的欧氏距离

$d[i, j] = \|X_i - X_j\|$, 则取最小聚类半径 $r_{\min} = d[i, j]_{\min}$, 最大聚类半径 $r_{\max} = d[i, j]_{\max}$, 在这一范围内选定一合适聚类半径 r 。定义一矢量 $A(l)$ 用于存放属于各类的输出矢量之和, $B(l)$ 用于存放属于各类的数目, 其中 l 表示类的数目。

Step2 假设利用减聚类算法已确定 M 个隐层结点即 M 个聚类中心, 分别为 c_1, c_2, \dots, c_M , 对于在线运行时采集到的第 k 个样本对 $X_k = (x_k, y_k)$ ($k = 1, \dots, N$), 计算 X_k 到 c_i ($i = 1, \dots, M$) 的距离 $\|X_k - c_i\|$, 然后找出最小距离 $d_{\min} = \|X_k - c_j\|$ 。如果 $d_{\min} < r$, 则 $A(j) = A(j) + y^k, B(j) = B(j) + 1$, 保持其它的 $A(i)$ 和 $B(i)$ 值不变, 此时隐层到输出层的权值为 $W(i) = A(i)/B(i)$ ($i = 1, \dots, M$)。如果 $d_{\min} > r$, 则增加一个新的聚类中心 $c_{M+1} = X_k, M = M + 1$, 保持 $A(i)$ 和 $B(i)$ 值不变, $A(M) = y_k, B(M) = 1$, 隐层到输出层的权值为 $W(M) = A(M)/B(M)$ 。

Step3 以高斯函数为基函数, 根据以上规则建立起来的 RBF 神经网络的输出为:

$$f(X_k) = \frac{\sum_{i=1}^M W_i \exp(-\frac{\|X_k - c_i\|^2}{r^2})}{\sum_{i=1}^M \exp(-\frac{\|X_k - c_i\|^2}{r^2})} \quad (6)$$

Step4 利用下式计算整个神经网络的误差性能指标:

$$E = \frac{1}{2} \sum_{i=1}^N (y_i - f(X_i))^2 \quad (7)$$

其中, X_i 为输入样本矢量, y_i 为输出样本, N 为样本数。当 E 小于阈值 E_m 时学习结束, 否则用 $r = r - h$ 修正聚类半径并返回第 2 步。

常规最近邻聚类算法当误差性能指标不达标时, 不论辨识误差多大, 聚类半径 r 均是递减一固定步长 h 。通过前面常规算法的描述可知, 聚类半径直接决定系统的辨识精度, 当辨识误差很大而步长又很小时, 系统学习的速度会较慢, 为解决这一问题, 聚类半径的调整根据误差大小采用可变步长, 当误差很大时, 加大步长, 加快学习速度, 当误差接近性能指标时, 取小步长, 避免增多类数导致过学习现象。式(8)就是一个可变步长的实例:

$$h = \begin{cases} 0.5, & e > 1 \\ 0.3, & 1 \geq e > 0.1 \\ 0.2, & 0.1 \geq e > 0.05 \\ 0.1, & 0.05 \geq e > 0.01 \end{cases} \quad (8)$$

3 仿真实例

为比较改进后的最近邻聚类算法与常规最近邻聚类算法对 RBF 神经网络训练速度的差别, 选择一数学

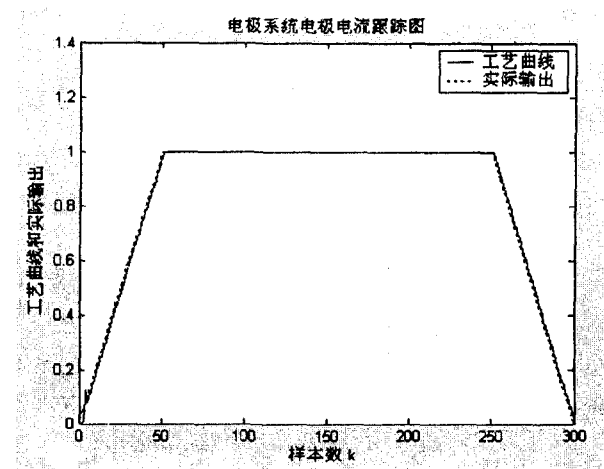
模型进行辨识, 数学模型如式(9)所示, 取样本 $n = 500$, 构造辨识模型为^[5]:

$$\hat{y}(k+1) = \hat{f}[u(k), u(k-1), u(k-2), y(k-1), y(k-2)],$$

$$\text{误差性能指标 } E = \frac{1}{2} \sum_{i=1}^N (y_i - \hat{f}(x_i))^2,$$

当误差性能指标相同时常规算法与改进算法的辨识结果是一样的, 如图 2 所示。两种算法的辨识参数比较见表 1, 从中可看出, 在误差指标相同的情况下改进算法使神经网络的学习次数明显减少, 这会大大加快网络的训练速度, 在实时性要求较高的非线性系统中是非常有效的。

$$\begin{cases} y(u(k)) = 0.6 \sin(\pi u(k)) + 0.2 \cos(3\pi u(k)) + \\ \quad 0.1 \sin(5\pi u(k)) \\ u(k) = \cos(k\pi/125) \end{cases} \quad (9)$$



(实线—实际输出, 虚线—辨识输出)

图 2 基于数学模型的辨识结果

表 1 常规算法与改进算法比较

	误差性能 指标 e	聚类半径 初值	聚类半径 终值	网络学习 次数 q	类数 M
常规 算法	0.05	$d_{\max} = 3.4691$	$d_{\min} = 0.1997$	344	167
改进 算法	0.05	$d_{\max} = 3.4691$	$d_{\min} = 0.1997$	75	167

4 应用实例

文中针对钢包精炼炉(LF 炉)电极系统这一非线性大滞后对象, 选择三相电极电流作为控制变量, 利用 RBF 神经网络通过现场高速采集的电流信号进行在线建模^[6], 基于该在线模型构造参考模型自适应控制器对电极系统进行实时控制。

图 3 为其结构图, 图中 TDL 表示将当前时刻的信

号进行若干次延迟, $i_m(k)$ 是 k 时刻电流参考模型的输出, $i(k)$ 是 k 时刻电极系统非线性对象的输出, $\hat{f}(k)$ 是 k 时刻辨识模型的输出。神经网络辨识器是对电极系统的辨识,其目的是利用当前及以前时刻对象的输入输出数据来预报下一步对象的输出应该是什么,预报准确度用 $e_i(k+1) = f(k+1) - \hat{f}(k+1)$ 来衡量。电流控制器的任务是借助神经网络来确定控制信号 $u(k)$,使得在相同输入情况下对象的输出 $i(k)$ 与参考模型 $i_m(k)$ (即期望输出) 之差不超过给定的误差范围。

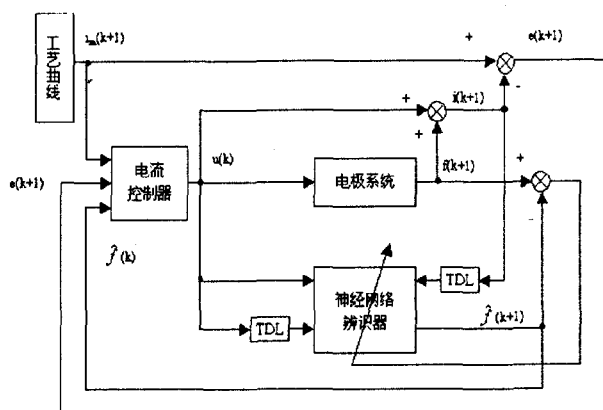


图3 基于RBF神经网络辨识的LF炉自适应控制系统结构图

将改进的最近邻聚类算法用于这一非线性系统的实时控制时,为进一步提高 RBF 神经网络的在线辨识速度,可以尝试限制在线辨识器的输入样本数,因为对于这种实时系统,若不限输入样本数,随着系统的不断运行,神经网络的规模会不断加大,系统辨识的速度会下降,从而达不到实时控制的目的。而实际上,对于这种大滞后非线性系统,系统下一时刻的输出只受当前时刻之前的一段时间内系统运行参数的影响,因此,限制在线辨识器的输入样本数是有一定意义的。图3中构造系统的在线辨识模型为:

$$\hat{i}(k+1) = \hat{f}[u_a(k), u_b(k), u_c(k), i(k), i(k-1), i(k-2), i(k-3), i(k-4)] \quad (10)$$

式(10)中, $u_b(k), u_c(k)$ 为三相电极系统中另两相电流控制器的输出,作为辨识器输入向量一部分的目的是考虑三相电极间的相互影响。

图4是本算法针对电极系统闭环运行时电极电流对梯形电流工艺曲线的跟踪结果。图中显示系统开始运行阶段误差较大,这主要是由于系统离线学习采集的样本不能真正反映系统的实际运行工况,运行一段时间后系统的在线学习能力开始体现出来,真正实现

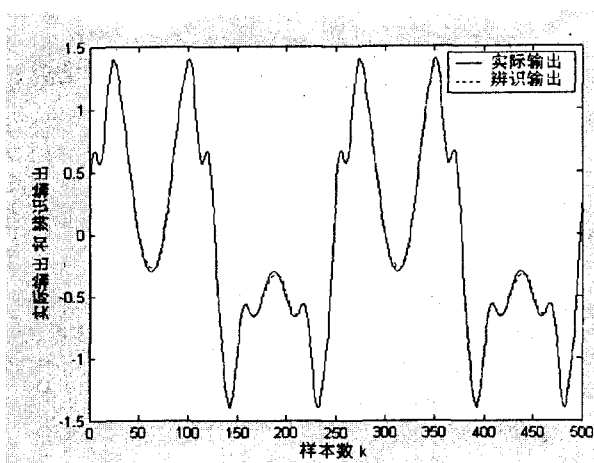


图4 电极系统电极电流跟踪图
了非线性实时系统的自适应控制。

5 结束语

根据常规最近邻聚类算法在实时系统预测中学习时间相对较长的问题,笔者提出了几个改进措施,将减聚类算法和最近邻聚类算法结合起来,使非线性系统的离线学习为在线学习提供了 RBF 神经网络的初始结构,系统在线学习时针对最近邻聚类算法采用了变步长策略。经函数拟合实验验证,这种改进算法明显缩短了 RBF 神经网络的学习时间。又将这种改进算法用到钢包精炼炉电极系统的在线辨识中,并对辨识样本进行限制,实际应用进一步表明文中对最近邻聚类算法的改进是有效的。

参考文献:

- [1] 储岳中,张绍德,张世峰.基于正则化 RBF 神经网络的钢包精炼炉电极系统智能建模[J].自动化与仪表,2004,19(5):5-7.
- [2] Chiu S L. Fuzzy Model Identification Based on Cluster Estimation[J]. Journal of Intelligent and Fuzzy System, 1994, 2(3): 756-759.
- [3] 王洪斌,杨香兰,王洪瑞.一种改进的 RBF 神经网络学习算法[J].系统工程与电子技术,2002,24(6):103-105.
- [4] 姚李孝,杨国清,王德意.减法聚类在水轮发电机模糊神经网络励磁调节器设计中的应用[J].电网技术,2004,28(16):46-49.
- [5] 王立新.模糊系统与模糊控制教程[M].北京:清华大学出版社,2003.
- [6] 储岳中,张绍德.基于 Visual C++ 的高速数据采集及其应用[J].工业控制计算机,2004,17(11):58-59.