第 15 卷第 2 期
1998 年 4 月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 15, No. 2
Apr., 1998

径向基函数神经网络的新型混合递推学习算法

汪小帆　王执铨　　　　宋文忠
(南京理工大学自动控制系・南京,210094)　(东南大学自动化所・南京,210096)

摘要：从径向基函数网络的硬件实现和实时应用的角度出发,给出了 RBF 网络的一种新型混合递推学习算法. 该算法既具有良好的数值性质又易于并行实现. 把 RBF 网络用于非线性系统在线辨识,仿真结果显示了本文方法的有效性.

关键词：径向基函数；神经网络；学习算法；系统辨识

1　引　　言

径向基函数神经网络(简称为 RBF 网络)是近年来受到模式识别、信号处理与控制等领域研究人员广泛关注的一种前馈神经网络模型[1]. 目前用于 BRF 网络的典型的递推算法包括,Moody 和 Darken 提出的混合自适应 n- 均值聚类和 LMS 算法[2];Chen 等人给出的混合聚类和 Givens 最小二乘法[3]. 在上述文献中,均采用 n- 均值聚类算法调整网络中心值,但该算法的一个缺点是,当隐节点选择不当时,会严重影响网络性能[4]. 此外,用 Givens 最小二乘法校正网络权值的优点在于,Givens 旋转具有良好的数值性质,且可用 Systolic 阵列并行实现,但 Givens 最小二乘法中的回代过程却是高度串行的,不易并行实现.

基于上述原因,我们提出了实时调整 RBFNN 中心值和权值的一种新的混合递推学习算法. 首先,通过对 n- 均值聚类算法的修正,给出了调整 RBFNN 中心值的一种新的自适应聚类算法,由于该算法可看作是文[4]中提出的次胜者受罚竞争学习(RPCL)算法的简化形式,固称其为 SRPCL 算法. 其次,基于逆 QR 分解[5],我们给出了校正 RBFNN 权值的一种新的递推学习算法——逆 Givens 最小二乘算法(简记为 IGLS 算法),该算法既保持了 Givens 旋转的优点,又避免了 Givens 最小二乘法中的回代步.

2　RBF 网络模型

RBF 网络是含有一个隐层的前馈神经网络. 为叙述方便,假设网络中有 M 个输入节点、H 个隐节点和 1 个输出节点,但本文所述方法可直接推广到有多个输出节点的情形.

设 RBF 网络的输入为 $V = [v_1, \cdots, v_M]^{\mathrm{T}}$, 则隐层节点的输出为

$$z_h = \Phi(\| V - C_h \|), \quad h = 1, 2 \cdots, H. \tag{1}$$

其中 $\| \cdot \|$ 通常取为欧氏范数, $C_h = [C_{h1}, \cdots, C_{hM}]^{\mathrm{T}} \in \mathbb{R}^M$ 称为网络的中心,基函数 Φ 有多种形式,本文中取为 $\Phi(x) = x^2 \log(x)$. 网络输出为

$$y = \sum_{h=1}^{H} W_h Z_h, \tag{2}$$

其中 w_h 称为网络的连接权.

3　混合 SRPCL-IGLS 算法

我们给出的 RBF 网络的混合递推学习算法包括两个子算法,即在每步迭代时作如下计算：i) 用一种新的自适应聚类算法——SRPCL 算法校正网络中心值；ii) 用一种新的递推最小

本文于 1996 年 1 月 9 日收到. 1997 年 4 月 22 日收到修改稿.

二乘算法——IGLS 算法校正网络权值.

3.1 SRPCL 算法

n- 均值聚类算法是一个经典的聚类算法. 但用此算法确定 RBF 网络中心值的主要问题是,网络中的隐节点数与输入模式的类别数必须相匹配,否则可能严重影响网络性能. 为此,我们将采用一种新的自适应聚类算法调整 RBFNN 的中心值. 该算法可看作是文[4]提出的次胜者受罚竞争学习(RPCL)算法的一种简化的自适应形式,我们称其为 SRPCL 算法. 其基本思想是,对每一个给定的输入,调整获胜单元的中心值,使之接近该输入,同时对次胜单元的中心值用较小的学习速率作反学习,使之偏离该输入.

给定初始值 $C_h(0), h=1,2,\cdots,H$,学习速率 α_c 和 $\alpha_r(0<\alpha_r\ll\alpha_c<1)$. 在输入第 t 个样本 $V(t)$ 时, SRPCL 算法按如下步骤调整网络中心值:

i) 计算
$$\beta_h = \| V(t) - C_h(t-1) \|, \quad h=1,\cdots,H.$$
$$\beta_k(t) = \min_{1\leqslant h\leqslant H} \beta_h(t), \quad \beta_l(t) = \min_{\substack{1\leqslant h\leqslant H\\ h\neq k}} \beta_h(t).$$

ii) 调整中心值
$$C_k(t) = C_k(t-1) + \alpha_c(V(t) - C_k(t-1)),$$
$$C_l(t) = C_l(t-1) - \alpha_r(V(t) - C_l(t-1)),$$
$$C_h(t) = C_h(t-1), \quad h\neq k, \quad h\neq l.$$

iii)
$$\beta_k(t) = \| V(t) - C_k(t) \|, \quad \beta_l(t) = \| V(t) - C_l(t) \|.$$

3.2 IGLS 算法

设时刻 t 为输入样本为 $V(t)=[v_1(t),\cdots,v_M(t)]^{\mathrm{T}}$, 隐层节点的输出为
$$z(t) = [z_1(t),\cdots,z_H(t)]^{\mathrm{T}} = [\Phi(\beta_1(t)),\cdots,\Phi(\beta_H(t))]^{\mathrm{T}}. \tag{3}$$
网络输出为 $y(t)$,记输出样本为 $d(t)$, 考虑如下的加权最小二乘问题:
$$\min_{w(t)} E(w(t)) = \| \Lambda(t)\bar{d}(t) - \Lambda(t)X(t)w(t) \|^2, \tag{4}$$
其中
$$w(t) = [w_1(t),\cdots,w_H(t)]^{\mathrm{T}}, \quad \bar{d}(t) = [d(1),\cdots,d(t)]^{\mathrm{T}},$$
$$\Lambda(1)=1, \quad \Lambda(t) = \begin{bmatrix} \sqrt{\lambda(t)}\Lambda(t-1) & 0 \\ 0 & 1 \end{bmatrix}, \quad X(t) = \begin{bmatrix} z^{\mathrm{T}}(1) \\ \vdots \\ z^{\mathrm{T}}(t) \end{bmatrix}, \tag{5}$$
$0<\lambda(t)\leqslant 1$ 为遗忘因子. 由矩阵理论知,存在正交矩阵 $Q(t)\in\mathbb{R}^{t\times t}$ 使得
$$Q(t)\Lambda(t)X(t) = \begin{bmatrix} R(t) \\ 0 \end{bmatrix}, \quad Q(t)\Lambda(t)\bar{d}(t) = \begin{bmatrix} q(t) \\ p(t) \end{bmatrix}. \tag{6}$$
其中 $R(t)$ 为 H 阶上三角矩阵,$q(t)$ 和 $p(t)$ 分别为 H 维和 $t-H$ 维向量. 于是有
$$\begin{aligned} E(w(t)) &= \| Q(t)\Lambda(t)\bar{d}(t) - Q(t)\Lambda(t)X(t)w(t) \|^2 \\ &= \left\| \begin{bmatrix} q(t) - R(t)w(t) \\ p(t) \end{bmatrix} \right\|^2. \end{aligned} \tag{7}$$
为使 $E(w(t))$ 极小,$w(t)$ 应满足方程
$$R(t)w(t) = q(t), \tag{8}$$
由回代过程可求得
$$w(t) = R_t^{-1}(t)q(t). \tag{9}$$
于是输入第 t 个样本时, 用 Givens 最小二乘(GLS)方法校正权的计算步骤为:

i）作 Givens 变换

$$G(t)\begin{bmatrix}\sqrt{\lambda(t)}R(t-1) & \sqrt{\lambda(t)}q(t-1)\\ Z^{\mathrm{T}}(t) & d(t)\end{bmatrix}=\begin{bmatrix}R(t) & q(t)\\ 0 & \zeta(t)\end{bmatrix}, \tag{10}$$

其中 $H+1$ 阶正交阵 $G(t)$ 为一系列 Givens 旋转阵之积.

ii）用逐步回代过程求解方程(8)得到权值 $w(t)$.

Givens 旋转具有良好的数值性质，且易于用 Systolic 阵列并行实现. 但求解方程(8)的回代过程却是高度串行的，不易并行实现. 这对 RBF 网络的硬件实现和实时应用是不利的. 为此，我们采用文献[5]的思想，基于逆 QR 分解，用逆 Givens 最小二乘法（IGLS 算法）校正 RBFNN 的权，其基本思想是，直接采用 Givens 变换修正 $R^{-\mathrm{T}}(t-1)$ 以得到 $R^{-\mathrm{T}}(t)$，从而免除了求解方程(8)的回代过程，使得 IGLS 算法既具有良好的数值性质，又易于并行实现. 下面给出具体的推导过程. 设

$$G(t)=\begin{bmatrix}A & b\\ g^{\mathrm{T}} & \delta\end{bmatrix}, \tag{11}$$

则由式(10)得

$$R(t)=\sqrt{\lambda(t)}AR(t-1)+bz^{\mathrm{T}}(t),\quad q(t)=\sqrt{\lambda(t)}Aq(t-1)+d(t)b, \tag{12}$$

把式(12)代入式(9)得

$$w(t)=[\sqrt{\lambda(t)}AR(t-1)+bz^{\mathrm{T}}(t)]^{-1}[\sqrt{\lambda(t)}Aq(t-1)+d(t)b], \tag{13}$$

由 $G(t)$ 的正交性可得

$$\begin{aligned} &b=-\frac{1}{\delta}Ag,\quad g=-\frac{\delta R^{-\mathrm{T}}(t-1)z(t)}{\sqrt{\lambda(t)}}=-\delta a(t),\\ &a(t)=\frac{R^{-\mathrm{T}}(t-1)z(t)}{\sqrt{\lambda(t)}}. \end{aligned} \tag{14}$$

再利用矩阵求逆引理

$$\begin{aligned} w(t)&=\left[\sqrt{\lambda(t)}R(t-1)+a(t)z^{\mathrm{T}}(t)\right]^{-1}\left[\sqrt{\lambda(t)}q(t-1)+d(t)a(t)\right]\\ &=\left[\frac{R^{-1}(t-1)}{\sqrt{\lambda(t)}}-\frac{R^{-1}(t-1)a(t)a^{\mathrm{T}}(t)}{\sqrt{\lambda(t)}(1+a^{\mathrm{T}}(t)a(t))}\right]\left[\sqrt{\lambda(t)}q(t-1)+d(t)a(t)\right]. \end{aligned} \tag{15}$$

定义

$$\gamma(t)=\sqrt{1+a^{\mathrm{T}}(t)a(t)},\quad u(t)=\frac{R^{-1}(t-1)a(t)}{\sqrt{\lambda(t)}\gamma(t)}, \tag{16}$$

最后得

$$w(t)=w(t-1)+\frac{u(t)}{\gamma(t)}[d(t)-z^{\mathrm{T}}(t)w(t-1)]. \tag{17}$$

可以证明[5]，存在一个 $H+1$ 阶正交矩阵 $P(t)$，使得

$$P(t)\begin{bmatrix}a(t) & \lambda(t)^{-1/2}R^{-\mathrm{T}}(t-1)\\ 1 & 0\end{bmatrix}=\begin{bmatrix}0 & R^{-\mathrm{T}}(t)\\ \gamma(t) & u^{\mathrm{T}}(t)\end{bmatrix}. \tag{18}$$

于是在输入第 t 个样本时，用 IGLS 算法校正 RBF 网络权的计算步骤为：

i）按式(14)计算 $a(t)$；

ii）按式(18)作 Givens 正交变换，得到 $\gamma(t)$，$u(t)$；

iii）按式(17)校正权.

记

$$R^{-\mathrm{T}}(t)=(r_{ij}(t))\in\mathscr{R}^{H\times H},$$

可以推得按 IGLS 算法校正权的显式递推格式如下:

$$\begin{cases} u_j^k(t) = 0, \quad 1 \leqslant j \leqslant H, \quad k < j. \\ \gamma^{(0)}(t) = 1. \end{cases} \tag{19}$$

$$\begin{cases} a_j(t) = \lambda(t)^{-½} \sum_{j=1}^{i} r_{ij}(t-1) z_j(t), \\ \gamma^{(i)}(t) = \sqrt{[\gamma^{(i-1)}(t)]^2 + a_i^2(t)}, \quad i = 1,2,\cdots,H. \\ s_i(t) = a_i(t)/\gamma^{(i)}(t), \\ c_i(t) = \gamma^{i-1}(t)/\gamma^{(i)}(t), \end{cases} \tag{20}$$

$$\begin{cases} r_{ij}(t) = \lambda(t)^{-½} C_i(t) r_{ij}(t-1) - s_i(t) u_j^{(i-1)}(t), \\ u_j^{(i)}(t) = c_i(t) u_j^{(i-1)}(t) + \lambda(t)^{-½} s_i(t) r_{ij}(t-1), \end{cases} \quad i = 1,2,\cdots,H; j = 1,\cdots,i. \tag{21}$$

$$w_h(t) = w_h(t-1) + \frac{u_h^{(H)}(t)}{\gamma^{(H)}(t)} \Big[d(t) - \sum_{i=1}^{H} W_i(t-1) z_i(t) \Big], \quad h = 1,2,\cdots,H. \tag{22}$$

4 RBF 网络用于非线性系统辨识

考虑如下的非线性系统

$$d(t+1) = f(d(t), d(t-1), d(t-2), u(t), u(t-1)), \tag{23}$$

其中 $u(t), d(t) \in \mathscr{R}$ 分别为系统的输入和输出,函数 f 具有如下形式

$$f(x_1, x_2, x_3, x_4, x_5) = \frac{x_1 x_2 x_3 x_5 (x_3 - 1) + x_4}{1 + x_3^2 + x_2^2}. \tag{24}$$

在辨识时采用串并结构

$$y(t+1) = N[d(t), d(t-1), d(t-2), u(t), u(t-1)], \tag{25}$$

其中 $y(t+1)$ 为网络在 $t+1$ 时刻的输出,$N[\cdot]$ 表示网络从输入到输出的函数关系. 取输入信号 $u(t)$ 为 $[-1,1]$ 上均匀分布的随机数,在不同的隐节点数下,分别采用文[4]中的混合 n- 均值聚类 ——-Ginens 最小二乘算法(简称算法 I)和本文给出的混合 SRPCKL-IGLS 算法(简称算法 Ⅱ)训练网络参数,在迭代 3000 次后,采用并行结构检验 RBF 网络的性能. 此时网络输出为

$$y(t+1) = N[y(t), y(t-1), y(t-2), u(t), u(t-1)]. \tag{26}$$

输入信号取为

$$u(t) = \sin(2\pi t/250), \quad t \leqslant 500;$$

$$u(t) = 0.8\sin(2\pi t/250) + 0.2\sin(2\pi t/25), \quad t > 500.$$

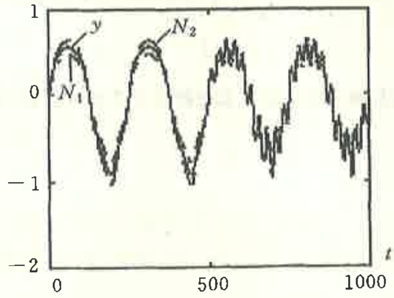

图 1 $H = 60$ 的辨识效果

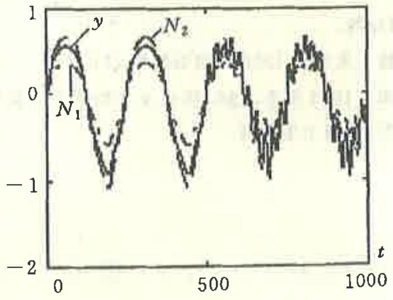

图 2 $H = 70$ 的辨识效果

仿真结果表明当 $50 \leqslant H \leqslant 60$ 时,采用算法 I 和算法 Ⅱ 均能得到满足的辨识效果(图 1),但算

法Ⅱ易于并行实现.而当 $65 \leqslant H \leqslant 75$ 时,算法Ⅱ的辨识效果明显优于算法Ⅰ(图2).

参考文献

1 Hush,D. R. and Hone,B. G.. Progress in supervised neural networks. IEEE Signal Processing Magazine,1993,1(1):8—37

2 Moody,J. and Darken,C. J.. Fast learning in networks of locally-tuned processing units. Neural Comput.,1989,2(2):281—284

3 Chen,S. et al.. Recursive hybrid algorithm for nonlinear systems identification using radial basis function networks. Int. J. Control,1992,55(5):1051—1070

4 Xu,L. et al.. Rival penalized competitive learning for clustering analysis,RBF net,and curve detection. IEEE Trans. Neural Networks,1993,4(4):636—649

5 Alexander,S. T. and Ghirniker,A. L.. A method for recursive least squares filtering based upon an Inverse QR decomposition. IEEE Trans. Signal Processing,1993,41(1):20—30

A New Hybrid Recursive Learning Algorithm for Radial Basis Function Neural Networks

WANG Xiaofan and WANG Zhiquan

(Department of Automatic Control,Nanjing University of Science & Technology · Nanjing,210094,PRC)

SONG Wenzhong

(Institute of Automation,Southeast University · Nanjing,210096,PRC)

Abstract: A new hybrid recursive learning algorithm for the radial basis function (RBF) neural network is proposed from the viewpoint of hardware implementations and on-line applications. The new algorithm has superior numerical properties and can be implemented in parallel easily. Finally,recursive identification of nonlinear systems using RBF network is investigated. The simulations show that the algorithm is very effective.

Key words: radial basis function; neual network; learning algorithm; system identification

本文作者简介

汪小帆 1967年生.1986年于苏州大数学系获学士学位,1991年于南京师范大学计算数学专业获硕士学位.1996年于东南大学自动化所获博士学位.现为南京理工大学自动控制博士后流动站研究人员,主要研究兴趣为非线性控制理论与应用、神经网络等.

王执铨 见本刊1998年第2期第271页.

宋文忠 1936年生.1960年毕业于南京工学院动力系,留校后长期从事工业自动化的教学和研究工作.现为东南大学自动化所教授,博士生导师.