

一种新型广义 RBF 神经网络在混沌时间序列预测中的研究*

李 军 刘君华

(西安交通大学电气工程学院, 西安 710049)

(2004 年 11 月 18 日收到 2005 年 3 月 14 日收到修改稿)

提出了一种新颖的广义径向基函数神经网络模型,其径向基函数(RBF)的形式由生成函数确定.然后,给出了易实现的梯度学习算法,同时为了进一步提高网络的收敛速度和网络性能,又给出了基于卡尔曼滤波的动态学习算法.为了验证网络的学习性能,采用基于卡尔曼滤波算法的新型广义 RBF 网络预测模型对 Mackey-Glass 混沌时间序列和 Henon 映射进行了仿真.结果表明,所提出的新型广义 RBF 神经网络模型能快速、精确地预测混沌时间序列,是研究复杂非线性动力系统辨识和控制的一种有效方法.

关键词:广义径向基函数神经网络,卡尔曼滤波,梯度下降学习算法,混沌时间序列,预测

PACC: 0545

1. 引 言

由确定性的非线性动力系统产生的复杂的类似随机现象的行为称为混沌现象.随着非线性混沌动力学的发展,混沌时间序列的预测已经成为一个非常重要的研究方向,并已在信号处理、自动控制等领域中得到了广泛的应用.混沌时间序列(或混沌信号)是指对一个混沌系统进行观测采样而得到的一个单变量时间序列.混沌时间序列的预测^[1]可看作动力学系统研究的“反问题”。“正问题”是给定非线性动力学系统,研究其相空间中轨道的各种性质.“反问题”是给定相空间中的一串迭代序列(轨道的演化过程)或一组观测序列.要构造一个非线性映射来表达原系统,这个映射就可作为预测模型.因此,如何构造预测模型是混沌时间序列预测中的一个关键问题.近年来,国内外学者利用各种适宜于复杂非线性系统建模的方法,如神经网络模型、模糊系统模型、模糊神经网络模型、支持向量机模型等,对混沌时间序列的预测进行了研究,并且取得了成功的应用^[2-6].例如,利用支持向量机模型预测混沌时间序列等,也是比较新颖的方法之一,但是为了保证其预测精度,也存在着如何选择核参数的困难以及计算时间的复杂性等问题.

径向基函数神经网络(RBFNN)作为前向神经网络的一种主要形式^[7,8],由于其拓扑结构简单,且具有显示表示学习过程的能力,学习快速、高效,使得其在需要实现实时自适应性能的系统(例如通信系统中)应用广泛^[9].已经从理论上证明,只要网络隐层的节点数足够多,RBF神经网络所执行的从 m 维输入空间到 n 维输出空间的非线性映射,可以以任意精度逼近任何单值连续函数^[10].因此,目前作为机器学习的方法也被广泛用于模式识别、函数逼近中.卡尔曼滤波器理论也被广泛用于前向多层神经网络和递归神经网络以及模糊系统的学习中^[11-13],其思想是将神经网络的训练看作对于具有过程噪声和测量噪声的非线性动力系统的状态最优滤波问题.

神经网络预测模型的成功与否很大程度上依赖于如何获得有效的学习算法及网络模型结构的选择.因此,本文首先提出一种新型广义 RBF 网络模型,其网络学习性能要优于常规的高斯型 RBF 网络,适合于多维输入-多维输出的一般情况.然后给出了其常规的梯度下降学习算法和基于卡尔曼滤波的动态快速学习算法.对具有时滞特性的 Mackey-Glass 混沌时间序列和 Henon 映射的实例仿真也验证了新型广义 RBF 网络模型的预测和辨识能力.基于卡尔曼滤波的广义新型 RBF 网络的预测模型不

* 国家自然科学基金(批准号:60276037)资助的课题.

仅在网络学习性能方面有显著改进,而且收敛速度更快,还便于实现在线自适应,因此,它是传统 RBF 神经网络方法的推广,可应用于多维输入、多维输出的一般非线性动力系统的辨识与实时控制,同时也是研究混沌时间序列预测的有效方法。

2. 新型广义 RBF 网络模型的建立

神经网络方法用于函数逼近是指,由有限数据样本中恢复其背后隐含的规律(函数映射),这是一个反问题,而且往往是不适定的(ill-posed)。解决这类问题的径向基函数神经网络(RBFNN)实质是完成正规化理论方法的逼近解^[7,8]。就函数逼近问题而言,正规化理论的基本思想是,在评价泛函中加入隐含某种先验信息(输入输出映射平滑性的约束)的辅助非负泛函项来稳定逼近问题的解。

本文将考虑多维输入、多维输出的一般情形下,广义 RBF 网络的结构,即对于映射 $x \rightarrow y, x \in R^q, y \in R^l, l > 1$,有

$$\hat{y}_i = w_{i,0} + \sum_{j=1}^c w_{i,j}g(\|x - v_j\|^2), 1 \leq i \leq l. \quad (1)$$

(1)式所代表的广义 RBF 网络具有 q 个输入, c 个径向基函数(RBF), l 个线性输出单元。如果 $g(x^2) = \phi(x)$, 并且 $\phi(\cdot)$ 是一个 RBF, 在给定一个训练样本集 $\{x_k, y_k\}_{k=1}^N, x \in R^q, y \in R^l$ 的情况下, 定义网络第 j 个隐节点的输出为:

$$h_{j,k} = \phi(\|x_k - v_j\|) = g(\|x_k - v_j\|^2), \quad 1 \leq j \leq c, 1 \leq k \leq N, \quad (2)$$

$$h_{0,k} = 1, 1 \leq k \leq N,$$

则广义 RBF 网络在输入矢量 x_k 下的输出为

$$\hat{y}_{i,k} = \sum_{j=0}^c w_{i,j}h_{j,k}, 1 \leq i \leq l, 1 \leq k \leq N. \quad (3)$$

基于此, 广义径向基函数 RBF 网络的一般网络拓扑结构如图 1 所示。

为了构造新类型的广义径向基函数, 即考虑作为新型广义 RBF 网络的隐节点函数 $\phi(x) = g(x^2)$ 的形式, 需首先满足下述公理化要求:

- (1) $\forall x_k \in \chi, v_j \in \nu, 0 < h_{j,k} < \infty$;
- (2) 如果 $\forall x_k, x_i \in \chi, v_j \in \nu, \|x_k - v_j\|^2 < \|x_i - v_j\|^2$, 那么 $h_{j,k} > h_{j,i}$;
- (3) 如果 $\forall x_k, x_i \in \chi, v_j \in \nu, \|x_k - v_j\|^2 < \|x_i - v_j\|^2$, 且定义梯度 $\nabla_{x_k} h_{j,k} \equiv \frac{\partial h_{j,k}}{\partial x_k}$, 那么

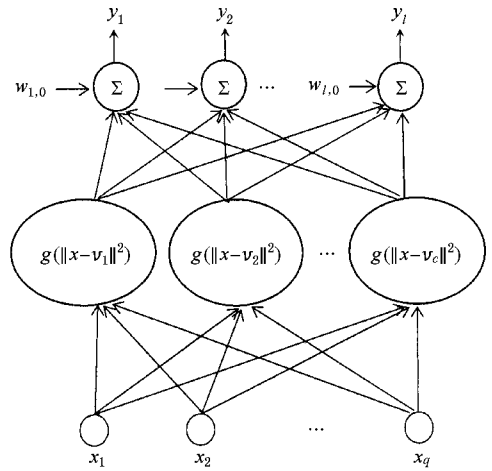


图 1 广义径向基函数网络结构

$$\frac{\|\nabla_{x_k} h_{j,k}\|^2}{\|x_k - v_j\|^2} > \frac{\|\nabla_{x_i} h_{j,i}\|^2}{\|x_i - v_j\|^2}.$$

因此, 按照这三个基本公理所选择的径向基函数(RBF)能推导出下面的定理^[14]。

定理 1 根据式 1 表示的 RBF 网络模型, 当且仅当 $g(\cdot)$ 是在 $(0, \infty)$ 上的连续函数时, 有下述性质:

- (1) $\forall x \in (0, \infty), g(x) > 0$;
- (2) $\forall x \in (0, \infty), g'(x) < 0$, 即 $g(x)$ 是一个在 $x \in (0, \infty)$ 的单调下降函数;
- (3) $\forall x \in (0, \infty), g''(x) > 0$, 即 $g'(x)$ 是一个在 $x \in (0, \infty)$ 的单调上升函数;
- (4) $\lim_{x \rightarrow 0} g(x) = L, L \in (0, \infty)$, 即 L 是一个有限数。

在此基础上, 若选择所允许的径向基函数的一般形式为 $\phi(x) = g(x^2)$, 且

$$g(x) = (g_0(x))^{\frac{1}{r-1}}, r \neq 1, g_0(x) \text{ 为生成函数,}$$

由定理 1 可推导出定理 2^[14], 它是新型广义径向基函数神经网络模型的重要理论依据。

定理 2 考虑式 (1) 表示的 RBF 网络模型, 当生成函数 $g_0(x)$ 是在 $(0, \infty)$ 上的连续函数时, 且 $g(x) = (g_0(x))^{\frac{1}{r-1}}, r \neq 1$. 若满足下面的条件时, 由 $g(x)$ 可以形成一个 RBF 网络。如果 $r > 1$ 则

- 1) $\forall x \in (0, \infty), g_0(x) > 0$ 且有 $\lim_{x \rightarrow 0^+} g_0(x) = L_1, L_1 \in (0, \infty)$;
- 2) $\forall x \in (0, \infty), g'_0(x) > 0$, 即 $g_0(x)$ 是一个在 $x \in (0, \infty)$ 的单调上升函数;
- 3) $\forall x \in (0, \infty), \frac{r}{r-1} (g'_0(x))^2 - g_0(x)g''_0(x)$

> 0.

如果 $r < 1$ 则

$$1) \forall x \in (0, \infty), g_0(x) > 0; \text{ 且有 } \lim_{x \rightarrow 0^+} g_0(x) =$$

$L_2, L_2 \in (0, \infty);$

2) $\forall x \in (0, \infty), g'_0(x) < 0$, 即 $g_0(x)$ 是一个在 $x \in (0, \infty)$ 的单调下降函数;

$$3) \forall x \in (0, \infty), \frac{r}{r-1} (g'_0(x))^2 - g_0(x) g''_0(x) < 0.$$

按照定理 2, 为了生成允许的 RBF 而形成新型广义 RBF 网络, 可选取指数型生成函数

$$g_0(x) = \exp(\beta x). \quad (4)$$

如果 $r > 1$, 令 $\sigma^2 = \frac{(r-1)}{\beta}$ 则有

$$\phi(x) = g(x^2) = \exp\left(-\frac{x^2}{\sigma^2}\right), \quad (5)$$

或选取线性生成函数,

$$g_0(x) = ax + b, a > 0, b \geq 0, \quad (6)$$

选取的 RBF 为 $g(x) = (g_0(x))^{-\frac{1}{r-1}}$, $r \neq 1$, 如果 $a = 1, r > 1, b = \gamma^2$ 则有

$$\phi(x) = g(x^2) = \left(\frac{1}{x^2 + \gamma^2}\right)^{\frac{1}{r-1}}. \quad (7)$$

若 $r = 3$, 所形成的特殊 RBF 即为逆多二次函数形式. RBF 网络的性能取决于网络的径向基函数(隐节点)的数目、节点中心、节点类型以及网络使用的有效学习算法. 本文以线性生成函数和指数型生成函数产生的新型 RBF 所构成的 RBF 网络为研究重点, 同时给出它们的梯度学习算法和基于动态卡尔曼滤波的快速学习算法.

3. 新型广义 RBF 网络的梯度学习算法

结合图 1 网络结构(q 维输入, l 维输出), 首先给出它的易实现的常规的梯度学习算法.

(1) 式可写成

$$\hat{y} = \begin{bmatrix} w_{1,0} & w_{1,1} & \dots & w_{1,c} \\ w_{2,0} & w_{2,1} & \dots & w_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ w_{l,0} & w_{l,1} & \dots & w_{l,c} \end{bmatrix} \begin{bmatrix} 1 \\ g(\|x - v_1\|^2) \\ \vdots \\ g(\|x - v_c\|^2) \end{bmatrix}. \quad (8)$$

若给定训练样本集为 $\{x_k, y_k\}_{k=1}^N$, 在输入矢量 x_k 作用下, 网络的输出为 \hat{y}_k , 则由(3)式可得

$$[\hat{y}_1 \dots \hat{y}_N] = \begin{bmatrix} w_{1,0} & w_{1,1} & \dots & w_{1,c} \\ w_{2,0} & w_{2,1} & \dots & w_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ w_{l,0} & w_{l,1} & \dots & w_{l,c} \end{bmatrix} \begin{bmatrix} 1 & \dots & \dots & 1 \\ g(\|x_1 - v_1\|^2) & \dots & \dots & g(\|x_N - v_1\|^2) \\ \vdots & \vdots & \ddots & \vdots \\ g(\|x_1 - v_c\|^2) & \dots & \dots & g(\|x_N - v_c\|^2) \end{bmatrix}. \quad (9)$$

由(1)式和(3)式可以定义权值矩阵

$$W = \begin{bmatrix} w_{1,0} & w_{1,1} & \dots & w_{1,c} \\ w_{2,0} & w_{2,1} & \dots & w_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ w_{l,0} & w_{l,1} & \dots & w_{l,c} \end{bmatrix} = \begin{bmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \vdots \\ \mathbf{w}_l^T \end{bmatrix}. \quad (10)$$

考虑(2)式及(9)式, 若定义网络隐层节点的输出矩阵为

$$H = [h_1 \dots h_N] = \begin{bmatrix} h_{0,1} & h_{0,2} & \dots & h_{0,N} \\ h_{1,1} & h_{1,2} & \dots & h_{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ h_{c,1} & h_{c,2} & \dots & h_{c,N} \end{bmatrix}, \quad (11)$$

则(9)式可写为矩阵形式

$$\hat{Y} = WH. \quad (12)$$

可用矩阵的 F -范数(Frobenius 范数)定义网络能量函数, 即

$$E = \frac{1}{2} \sum_{k=1}^N \sum_{i=1}^l (y_{i,k} - \hat{y}_{i,k})^2 = \frac{1}{2} \|Y - \hat{Y}\|_F^2. \quad (13)$$

使用梯度下降法, 网络权值向量的调整量为

$$\Delta w_i = -\eta \nabla_{w_i} E = \eta \sum_{k=1}^N \epsilon_{i,k}^o \mathbf{h}_k \quad (i = 1, \dots, l), \quad (14)$$

其中, η 是学习率, 输出层误差 $\epsilon_{i,k}^o = y_{i,k} - \hat{y}_{i,k}$.

网络隐层 RBF 的中心矢量的调整量为

$$\Delta v_j = -\eta \nabla_{v_j} E = \eta \sum_{k=1}^N (\epsilon_{j,k}^h) (x_k - v_j) \quad (j = 1, \dots, c), \quad (15)$$

其中, 网络隐层误差

$$\epsilon_{j,k}^h = -2g'(\|x_k - v_j\|^2) \sum_{i=1}^l (\epsilon_{i,k}^o) w_{i,j}.$$

由于 $g(x) = (g_0(x))^{1-m}$ 则

$$g'(x) = \frac{1}{1-m} (g(x))^m g_0(x).$$

所以有：

$$\epsilon_{j,k}^h = \frac{2}{m-1} (h_{j,k})^m g'_0(\|x_k - v_j\|^2) \sum_{i=1}^l (\epsilon_{i,k}^o) w_{i,j}. \tag{16}$$

按(6)式选取线性生成函数 $g_0(x) = ax + b$,代入(16)式,并考虑(2)式和(7)式得

$$\epsilon_{j,k}^h = \frac{2}{m-1} (\|x_k - v_j\|^2 + \gamma^2)^{\frac{m}{1-m}} \sum_{i=1}^l (\epsilon_{i,k}^o) w_{i,j}. \tag{17}$$

按(4)式选取指数型生成函数 $g_0(x) = \exp(\beta x)$,代入式(16),并考虑(2)式和(5)式得

$$\epsilon_{j,k}^h = \frac{2}{\sigma^2} \exp\left(-\frac{\|x_k - v_j\|^2}{\sigma^2}\right) \sum_{i=1}^l (\epsilon_{i,k}^o) w_{i,j}. \tag{18}$$

令 s 为迭代次数,新型 RBF 网络的网络权值矢量和隐层 RBF 中心矢量的梯度下降算法更新公式为：

$$\begin{aligned} w_i(s+1) &= w_i(s) + \Delta w_i \quad (i = 1 \dots l), \\ v_j(s+1) &= v_j(s) + \Delta v_j \quad (j = 1 \dots c). \end{aligned} \tag{19}$$

基于同样的梯度下降算法,指数型生成函数和线性生成函数所形成的新型广义 RBF 网络均比常规的高斯性 RBF 网络性能有很好的改善^[14].而且,线性生成函数所形成的新型广义 RBF 网络无论在测试集上还是在训练集上,网络性能都更好.例如,当 r 的数值选取为 2—4,且 b 的数值接近 1 时,经训练后的该 RBF 网络性能更佳.

4. 新型广义 RBF 网络的卡尔曼滤波快速学习算法

为了进一步提高新型 RBF 网络模型的性能,结合线性动态系统卡尔曼滤波的思想,将卡尔曼滤波用于多维输入—多维输出的新 RBF 网络参数的优化上,将具有不同于其常规梯度下降算法的快速收敛速度.实质是将网络的训练问题看作一个非线性动力系统的动态参数估计问题来处理^[11—13],即随时序对参数不断进行重估,使得网络的输出与理想输出之间的误差随时序的增加而逐渐减小.

4.1. 非线性离散时间系统的卡尔曼滤波方程

首先考虑下面的有限维离散非线性时间系统的状态方程和观测方程：

$$\begin{cases} \theta_{k+1} = f(\theta_k) + \omega_k, \\ y_k = h(\theta_k) + \xi_k, \end{cases} \tag{20}$$

其中 θ_k 是状态矢量, y_k 为观测矢量, ω_k 是过程噪声, ξ_k 是量测噪声.

卡尔曼滤波就是在给定样本的量测值 y_k 下,求出状态 θ_k 的估计值 $\hat{\theta}_k$.假定初始状态为 θ_0 , ω_k 和 ξ_k 均为零均值白噪声,其统计特性如下：

$$\begin{aligned} E(\theta_0) &= \bar{\theta}_0, \\ E(\omega_k) &= E(\xi_k) = 0, \\ E(\omega_k \omega_j^T) &= Q \delta_{k,j}, E(\xi_k \xi_j^T) = R \delta_{k,j}, \tag{21} \\ E[(\theta - \theta_0)(\theta - \theta_0)^T] &= P_0, \tag{22} \end{aligned}$$

其中 E 为数学期望符号,而 $\delta_{k,j} = \begin{cases} 1 & k=j \\ 0 & k \neq j \end{cases}$.将(20)式中的非线性函数项围绕滤波估计值 $\hat{\theta}_k$ 进行 Taylor 级数展开,略去二次以上高次项,则有：

$$\begin{cases} f(\theta_k) = f(\hat{\theta}_k) + \left. \frac{\partial f(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}_k} \cdot (\theta_k - \hat{\theta}_k), \\ h(\theta_k) = h(\hat{\theta}_k) + \left. \frac{\partial h(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}_k} \cdot (\theta_k - \hat{\theta}_k). \end{cases} \tag{23}$$

定义 $F_k = \left. \frac{\partial f(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}_k}$, $H_k^T = \left. \frac{\partial h(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}_k}$ (20)式可表示为

$$\begin{cases} \theta_{k+1} = F_k \theta_k + \omega_k + \psi_k, \\ y_k = H_k^T \theta_k + \xi_k + j_k, \end{cases} \tag{24}$$

其中 $\psi_k = f(\hat{\theta}_k) - F_k \hat{\theta}_k$, $j_k = h(\hat{\theta}_k) - H_k^T \hat{\theta}_k$.同基本卡尔曼滤波方程的模型相比,只增加了非随机的外作用项,因此,可得到递推的最佳估计 $\hat{\theta}_k$

$$\begin{cases} \hat{\theta}_k = f(\hat{\theta}_{k-1}) + K_k [y_k - h(\hat{\theta}_{k-1})] \\ K_k = P'_k H_k (R + H_k^T P'_k H_k)^{-1}, \\ P'_k = F_k P_{k-1} F_k^T + Q, \\ P_k = (I - K_k H_k^T) P'_k, \end{cases} \tag{25}$$

其中 K_k 为卡尔曼增益矩阵, P'_k 是一步预测误差的均方阵, P_k 是状态估计的均方误差矩阵,即状态估计误差的协方差阵.

对于线性系统而言,卡尔曼滤波的状态估计本质上是最优滤波器,达到了条件均值,即 $\hat{\theta}_{k+1} =$

$E[\theta_{k+1} | (y_0, y_1, \dots, y_k)]$. 对于非线性系统, 由于采取了近似线性化处理, 这种滤波估计是次优的, 接近其条件均值.

4.2. 新型广义 RBF 网络的快速卡尔曼滤波算法

结合图 1 所示的 q 维输入, l 维输出的新型广义 RBF 网络, 网络隐层 RBF 的中心矢量为 $\nu_j (1 \leq j \leq c)$, 训练样本集为 $\{x_k, y_k\}_{k=1}^N$, 误差矢量就是 RBF 网络的输出与理想目标值的差. 网络输出的目标矢量为 y , $h(\hat{\theta}_k)$ 为网络优化过程中第 k 次迭代的实际输出. 定义

$$y = [y_{1,1} \dots y_{1,N} \dots y_{l,1} \dots y_{l,N}]^T,$$

$$h(\hat{\theta}_k) = [\hat{y}_{1,1} \dots \hat{y}_{1,N} \dots \hat{y}_{l,1} \dots \hat{y}_{l,N}]^T. \quad (26)$$

为应用(25)式的卡尔曼滤波动态递推算算法, 将新型 RBF 网络的实际输出看作非线性系统的输出, 网络权值矢量和隐层节点中心矢量均看作是非线性系统的状态向量, 并考虑人为加入噪声, 则有:

$$\begin{cases} \theta_{k+1} = \theta_k + \omega_k, \\ y_k = h(\theta_k) + \xi_k, \end{cases} \quad (27)$$

其中 $\theta = [w_1^T \dots w_l^T v_1^T \dots v_c^T]^T$. 系统状态向量 θ 的元素共有 $(l(c+1) + qc)$ 个.

首先确定 H_k , 由(26)式和(3)式可得

$$\begin{aligned} \hat{Y} &= h(\hat{\theta}) = [\hat{y}_{1,1} \dots \hat{y}_{1,N} \dots \hat{y}_{l,1} \dots \hat{y}_{l,N}]^T \\ &= \left[\sum_{i=0}^c w_{1,i} h_{i,1} \dots \sum_{i=0}^c w_{1,i} h_{i,N} \right. \\ &\quad \left. \dots \sum_{i=0}^c w_{l,i} h_{i,1} \dots \sum_{i=0}^c w_{l,i} h_{i,N} \right]^T \end{aligned}$$

$$H_v = \frac{\partial \hat{Y}^T}{\partial v} = \begin{bmatrix} \frac{\partial v_1}{\partial v} \\ \frac{\partial v_c}{\partial v} \end{bmatrix} \left[w_{1,0} + \sum_{j=1}^c w_{1,j} g(\|x_1 - v_j\|^2) \dots w_{1,0} + \sum_{j=1}^c w_{1,j} g(\|x_N - v_j\|^2) \dots \right.$$

$$w_{l,0} + \sum_{j=1}^c w_{l,j} g(\|x_1 - v_j\|^2) \dots w_{l,0} + \sum_{j=1}^c w_{l,j} g(\|x_N - v_j\|^2) \left. \right]$$

$$= \begin{bmatrix} -w_{1,1} g'_{1,1} \mathcal{A}(x_1 - v_1) & \dots & -w_{1,1} g'_{1,N} \mathcal{A}(x_N - v_1) & \dots & -w_{1,1} g'_{1,1} \mathcal{A}(x_1 - v_1) & \dots & -w_{1,1} g'_{1,N} \mathcal{A}(x_N - v_1) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -w_{1,c} g'_{1,c} \mathcal{A}(x_1 - v_c) & \dots & -w_{1,c} g'_{1,N} \mathcal{A}(x_N - v_c) & \dots & -w_{1,c} g'_{1,c} \mathcal{A}(x_1 - v_c) & \dots & -w_{1,c} g'_{1,N} \mathcal{A}(x_N - v_c) \end{bmatrix}, \quad (30)$$

其中 $g'_{i,j} = g'(\|x_i - v_j\|^2)$ 若 $g_0(x) = x + \gamma^2$ 则

由 $g'(x) = \frac{1}{1-r}(g(x))g_0(x)$ 及(2)式 则

$$= \left[w_{1,0} + \sum_{j=1}^c w_{1,j} g(\|x_1 - v_j\|^2) \right.$$

$$\dots w_{1,0} + \sum_{j=1}^c w_{1,j} g(\|x_N - v_j\|^2) \left. \dots w_{l,0} + \sum_{j=1}^c w_{l,j} g(\|x_1 - v_j\|^2) \right.$$

$$\dots w_{l,0} + \sum_{j=1}^c w_{l,j} g(\|x_N - v_j\|^2) \left. \right]^T. \quad (28)$$

若定义 $w = [w_1^T \dots w_l^T]^T$ 则有

$$H_w = \frac{\partial \hat{Y}^T}{\partial w} = \begin{bmatrix} \frac{\partial w_{1,0}}{\partial w} \\ \frac{\partial w_{1,c}}{\partial w} \\ \frac{\partial w_{l,0}}{\partial w} \\ \frac{\partial w_{l,c}}{\partial w} \end{bmatrix}$$

$$\times \left[\sum_{i=0}^c w_{1,i} h_{i,1} \dots \sum_{i=0}^c w_{1,i} h_{i,N} \right.$$

$$\left. \dots \sum_{i=0}^c w_{l,i} h_{i,1} \dots \sum_{i=0}^c w_{l,i} h_{i,N} \right]$$

$$= \begin{bmatrix} H & 0 & \dots & 0 \\ 0 & H & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & H \end{bmatrix}, \quad (29)$$

其中 H 由(11)式定义, 它是一个 $(c+1) \times N$ 矩阵, 所以 H_w 是一个 $l \times [(c+1) + N]$ 矩阵.

若定义 $v = [v_1^T \dots v_c^T]^T$ 则有

$$g'_{i,j} = g'(\|x_i - v_j\|^2) = \frac{1}{1-r} h'_{i,j}. \quad (31)$$

若 $g_0(x) = \exp(\beta x)$, 则由 $g'(x) = \frac{1}{1-r}(g(x))$

$g_0(x)$ 及 (2) 式, 则

$$g'_{i,j} = g'(\|x_i - v_j\|^2) = \frac{\beta}{1-r} h_{j,i}, \quad (32)$$

可看出, H_v 是一个 $qc \times lN$ 矩阵. 最后, 由 θ 及 H_k^T 的定义, 可知

$$H_k = \begin{bmatrix} H_w \\ H_v \end{bmatrix}. \quad (33)$$

F_k 的确定很容易, 对照 (24) 式与 (27) 式, 有 $f(\theta) = \theta$, 此时 $F_k = I$ (单位阵).

Q 和 R 是人为加入的系统过程噪声 ω_k 、量测噪声 ξ_k 各自的协方差阵, 均可以在初始化时给定. 状态估计的均方误差矩阵初值 P_0 也在初始化时给定. 虽然 y_k 实际上是 RBF 网络的输出目标常值矢量, 但被写成迭代次数 k 的函数, 当 F_k, H_k 确定后以及 Q, R 和 P_0 给定后, 就可以应用 (25) 式的递推卡尔曼滤波算法求得状态 θ_k 的最终估计值 $\hat{\theta}_k$, 算法收敛的条件是 (13) 式的网络误差函数低于 0.001 时迭代结束. 在以下的实例仿真时, P_0, Q 和 R 都在初始化时给定一个较大的数值, 如 $P_0 = 40I, Q = 40I, R = 40I$. 其中 I 为 $[(c+1)+qc] \times [(c+1)+qc]$ 的单位阵, 系统的状态矢量 θ_0 为 0 矢量, 即网络的权值矢量和其隐层的中心矢量均初始化为 0.

5. 新型 RBF 网络在混沌时间序列中的应用

混沌时间序列的一个显著特征是对初始状态极其敏感, 这给混沌时间序列的预测带来很大的难度. 研究表明, 根据混沌动力系统的相空间延迟坐标重构理论, 基于神经网络的强大非线性映射能力, 就可建立混沌时间序列的预测模型^[2,8].

若动力系统由下述离散时间非线性差分方程表示

$$x(n+1) = F(x(n)), \quad (34)$$

其中 $x(n)$ 是系统在 n 时刻的状态向量, $F(\cdot)$ 是非线性向量函数, 则观测到的时间序列 $y(n)$ 为

$$y(n) = g(x(n)) + m(n), \quad (35)$$

其中 $m(n)$ 为噪声, $g(\cdot)$ 是标量函数. 根据 Takens 定理^[15], 原动力学系统运动轨道的几何结构可由下述不带噪声的 $y(n)$ 在 D 维空间“重构”

$$y_R(n) = [y(n), y(n-1), \dots, y(n-(D-1)\tau)]^T, \quad (36)$$

其中 τ 称为归一化嵌入时延宽度. 重构的充分条件是 $D \geq 2d + 1$, d 是动力系统的关联维数. 重构的结果可视为原系统相应的一条轨道在 R^D 中的展开 (或“嵌入”). 也可用下述 q 维向量作为重建向量, 即

$$y_R(n) = [y(n), y(n-1), \dots, y(n-(q-1))]^T, \quad (37)$$

其中, q 为整数, 且 $q \geq D_E \tau$, D_E 为嵌入维数. 为了预测, 要解决两个问题, 一是构造一个用于嵌入的短时记忆结构, 用于产生 $y_R(n)$, 二是构造一个多输入单输出的非线性系统模型 $f: R^q \rightarrow R$ 可作为一步预测用. 选择合适的神经网络模型, 就可以完成这种预测. 模型的参数由样本学习, 即

$$\hat{y}(n+1) = f(y_R(n)). \quad (38)$$

因此应用本文提出的新型广义 RBF 网络作为预测模型即是一个选择, 初始化时首先应选择一个合适的输入向量维数 q 作为重构向量, 并自适应地选择新型广义 RBF 网络的隐层中心节点数目, 一种可行的方法是交叉验证 (cross-validation), 自适应地调整网络隐层中心节点 c 的大小.

网络预测模型所需的输入向量为 $X(N \times q)$ 和输出向量为 $Y(N \times l)$, N 为时间序列的总长; 为了保证网络训练的精度, 需对网络样本数据进行归一化预处理, 即要对输入向量 X 和输出向量 Y 进行预处理

$$\begin{cases} X(k, i) = \frac{X(k, i) - \text{mean}(i)}{\text{std}(i)} \\ Y(k, j) = \frac{Y(k, j) - \text{mean}(j)}{\text{std}(j)} \end{cases}, \quad (39)$$

$k = 1, 2, \dots, N; i = 1, 2, \dots, q; j = 1, 2, \dots, l,$

其中 $\text{mean}(i), \text{std}(i)$ 分别是输入向量 X 的第 i 列的算术平均值和标准方差; mean_j 和 std_j 分别是输出向量 Y 的第 j 列的算术平均值和标准方差.

下面的混沌时间序列仿真实例中, 新型广义 RBF 网络的 RBF 将选取性能更好的线性生成函数 $g_0(x) = x + \gamma^2$, 其中 $\gamma = 1, m = 3, c = 15$, 这些参数也可以根据交叉验证的方法调整. 时间序列的预测输出 Y 是一维向量, 即 $l = 1$. 由于卡尔曼滤波算法相比梯度学习算法的收敛速度要快得多, 而且网络的性能很好, 因此, 以下的实例中就仅给出基于动态卡尔曼滤波快速算法的仿真结果. 需要注意得是, 如用梯度学习算法, 训练时的学习率可选小些, 如选取

$$\eta = 10^{-4}.$$

另外,采用均方根误差(mean square error, MSE)

$$MSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (y(j) - \hat{y}(j))^2} \quad (40)$$

作为评价模型整体预测效果的指标.其中 $y(j)$ 和 $\hat{y}(j)$ 分别为混沌时间序列的实际值和预测值.

5.1. Mackey-Glass 的混沌时间序列预测

自 Mackey 和 Glass 首次发现时滞系统中的混沌现象以来^[6],时滞混沌系统便引起了人们广泛关注,并常常用来检验非线性系统模型的性能,其混沌时间序列可由下面的时滞差分方程产生:

$$\frac{dx(n)}{dn} = \frac{0.2x(n-\tau)}{1+x^{10}(n-\tau)} - 0.1x(n), \quad (41)$$

其中 τ 为时滞参数, Mackey-Glass 方程的行为特性已在文献^[17]做过深入研究. $\tau > 17$ 时呈现混沌性. τ 值越大,混沌程度越高.选取 $\tau = 17$,为了获取每一积分时间点的时间序列值,用四阶龙格-库塔法寻找方程(41)的数值解,时间步长为 0.1,这样经过对数值积分,就得到了时间序列 $x(n)$.取 $q = 4$,抽取 2500 对输入输出数据,即 $[x(t-4), x(t-3), x(t-2), x(t-1), x(t)]$,其中 $t = 4$ 到 2503.前 2000 个数据对用于训练集,后 1000 个数据对验证模型的准确性.用基于快速卡尔曼滤波算法的新型广义 RBF 网络进行训练建模,迭代 12 次后,网络收敛.图 2 表示了从 $t = 4$ 到 $t = 2503$ 的系统的实际输出与所建模型的预测值比较曲线,模型各点的预测误差 Error 分布见图 3,整个训练集和测试集的 MSE 误差为 0.00641.从图 2 中不难看出,新型 RBF 网络模型的预测值与系统的实际值输出吻合得很好.

5.2. Henon 映射

Henon 映射是最著名的简单动力学系统之一,其动力学方程为:

$$\begin{cases} x(k+1) = 1 + y(k) - ax^2(k), \\ y(k+1) = bx(k), \end{cases} \quad (42)$$

其中,当 $a = 1.4, b = 0.3$ 时,系统成为混沌态^[18].利用(42)式的迭代计算,可得到时间序列.抽取 500 对输入输出数据,即 $[x(t-4), x(t-3), x(t-2), x(t-1), x(t)]$.前 400 个数据对用于训练集,后 100 个数据对验证模型的准确性.用基于快速卡尔曼滤波算法的新型广义 RBF 网络进行训练建模,迭代 14 次后,网络收敛.图 4 表示了 Henon 时间序列的实际

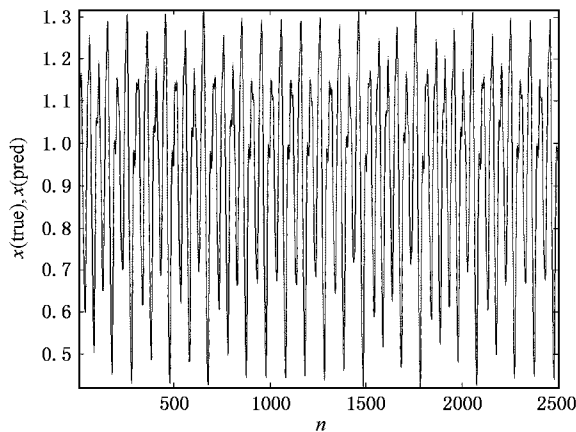


图 2 Mackey-Glass 时间序列实际值(实线)和预测值(点线)

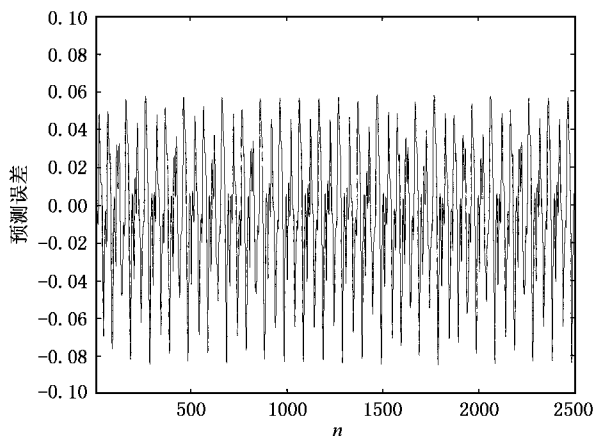


图 3 Mackey-Glass 模型的预测误差(Error 分布)

输出与所建模型的预测值比较曲线,模型各点的预测误差 Error 分布见图 5,其 MSE 误差为 0.0413.从图 4 中不难看出,新型 RBF 网络模型的预测值与系统的实际值输出吻合得很好.而根据预测模型重建

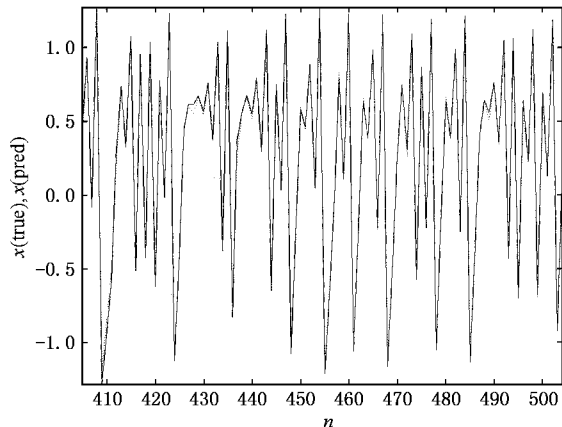


图 4 Henon 混沌时间序列的实际值(实线)和模型的预测值(虚线)

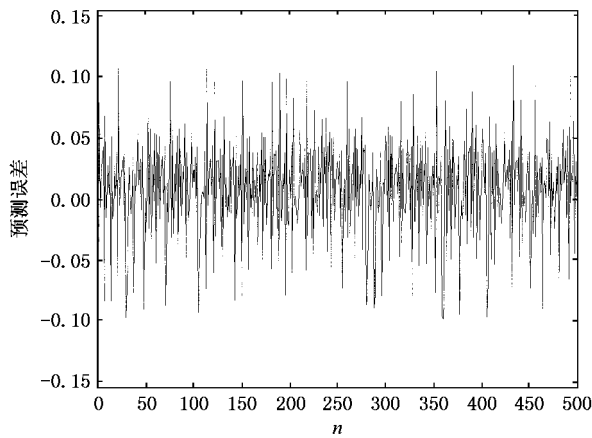


图5 Henon 模型的预测误差(Error 分布)

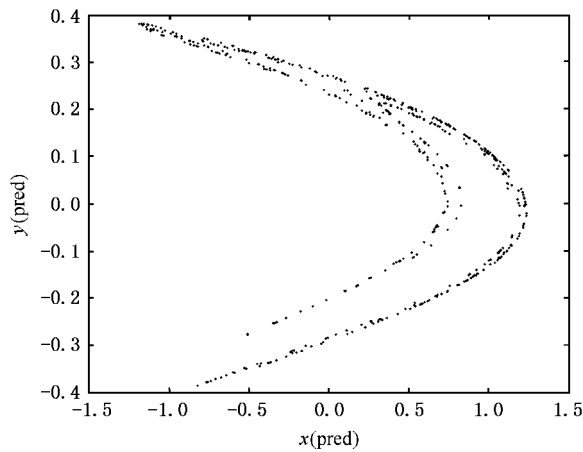


图6 Henon 预测模型的吸引子

的吸引子如图 6 所示,与原系统的吸引子基本一致,可见预测模型较好复现了系统的动态特性.

6. 结 论

本文提出一种适合多输入多输出的一般情形下的新型广义 RBF 网络模型并用于混沌时间序列预测,其性能优于常规的高斯径向基函数神经网络.同时给出了易实现的梯度学习算法和结合卡尔曼滤波理论的快速动态算法.对 Mackey-Glass 时间序列和 Henon 映射的仿真实例很好地验证了基于卡尔曼滤波学习算法的新型广义 RBF 网络预测模型之性能的有效性.它的网络学习速度很快,克服了普通梯度算法中权值初始化不好时易对网络收敛性能有影响的不足.因此,本文提出的基于新型广义 RBF 神经网络模型的方法是传统神经网络(包括 RBF 网络)方法用于混沌时间序列预测模型的推广,为混沌时间序列预测和复杂系统建模、辨识提供了一个极有吸引力的新途径.

- [1] Casdagli M 1989 *Physica D* **35** 335
- [2] Ramazan G 1997 *Physica D* **108** 119
- [3] Chen L and Chen G R 2000 *IEEE Trans. Circuits and Systems I* **47** 1527
- [4] Tan W, Wang Y N, Zhou S W *et al* 2003 *Acta Phys. Sin.* **52** 795 (in Chinese) [谭文、王耀南、周少武等 2003 物理学报 **52** 795]
- [5] Cui W Z, Zhu C C, Bao W X *et al* 2004 *Acta Phys. Sin.* **53** 3303 (in Chinese) [崔万照、朱长纯、保文星等 2004 物理学报 **53** 3303]
- [6] Thissen U, Brakel R, Weijer A P *et al* 2003 *Chemometrics and Intelligent Laboratory Systems* **69** 35
- [7] Poggio T and Girosi F 1990 *Science* **247** 978
- [8] Haykin S 2001 *Neural Networks: A Comprehensive Foundation* (New Jersey: Prentice-Hall)
- [9] Mulgrew B 1996 *IEEE Signal Processing Magazine* **13** 50
- [10] Park J and Sandberg I W 1993 *Neural Computation* **5** 305
- [11] Puskorius G V and Feldkamp L A 1994 *IEEE Trans. Neural Networks* **5** 279
- [12] Isabelle R and Personnag L 1998 *Neurocomputing* **20** 279
- [13] Simon D 2003 *Fuzzy Sets and Systems* **132** 189
- [14] Karayiannis N B and Randolph-Gips M M 2003 *IEEE Trans. Neural Networks* **14** 835
- [15] Takens F 1981 *Dynamical Systems and Turbulence* **898** 366
- [16] Mackey M C and Glass L 1977 *Science* **197** 287
- [17] Farmer J D 1982 *Physica D* **3** 366
- [18] Henon M 1976 *Commun. Math. Phys.* **50** 69

On the prediction of chaotic time series using a new generalized radial basis function neural networks^{*}

Li Jun Liu Jun-Hua

(School of Electronical Engineering , Xi 'an Jiaotong University , Xi 'an 710049 , China)

(Received 18 November 2004 ; revised manuscript received 14 March 2005)

Abstract

Radial basis function (RBF) networks have been widely used for function approximation and pattern classification as an alternative to conventional feedforward neural networks. A novel generalized RBF neural network model is presented. The form of RBF is determined by a generator function , and then an easily implementable gradient decent learning algorithm for training the new generalized RBF networks is given. Simultaneously , a fast dynamic learning algorithm based on Kalman filter is also proposed to improve the performance and accelerate the convergence speed of the new generalized RBF networks . The generalized RBF neural networks based on Kalman filtering dynamic learning algorithm is then applied to the chaotic time series prediction on the Mackey-Glass equation and the Henon map to test the validity of this proposed model. Simulation results show that the new generalized RBF networks can accurately predict chaotic time series. It provides an attractive approach to study the properties of complex nonlinear system model and chaotic time series .

Keywords : generalized radial basis function neural networks , Kalman filter , gradient decent learning algorithm , chaotic time series prediction

PACC : 0545

^{*} Project supported by the National Natural Science Foundation of China (Grant No. 60276037)