第 26 卷第 4 期
1997 年 8 月

信 息 与 控 制
Information and Control

Vol. 26, No. 4
Aug., 1997

RBF 神经网络理论及其在控制中的应用[*]

王旭东　邵惠鹤
(上海交通大学自动化研究所　上海　200030)

摘　要　对 RBF 神经网络的结构、分类、函数逼近理论及训练方法进行了综述, 并且对 RBF 网络的优点及问题作了分析, 同时介绍了目前 RBF 网络在控制方面主要应用情况, 最后提出了 RBF 网络在控制中的研究及应用新方向.

关键词　RBF 神经网络, 全局逼近, 最佳逼近, 预测控制, 内模控制, 软测量, 操作优化

1　前言

人工神经元网络是一门交叉学科, 在许多方面得到了应用. 从神经网络的基本模式看, 主要有: 前馈型、反馈型、自组织型及随机型网络. 这 4 种类型各自具有不同的网络模型. 在前馈网络中主要有 Adaline, BP 网及 RBF 网络; 反馈网络主要有 Hopfield 网络; 自组织网络主要有 ART 网; 随机网络主要有 Boltzman 机. 最近由于模糊及分形与人工神经网络的结合形成了模糊神经网络和分形神经网络.

目前, 在控制领域内神经网络正在稳步地发展, 这种发展的动力主要来自 3 个方面[1]: (1) 处理越来越复杂的系统的需要; (2) 实现越来越高的设计目标的需要; (3) 在越来越不确定的情况下进行控制的需要.

在控制中, 应用较多的网络是 BP 网络, 但 BP 网络存在局部最优问题, 并且训练速度慢, 效率低. RBF 网络在一定程度上克服了这些问题, 因此它的研究与应用越来越得到重视. 本文综述了 RBF 神经网络的有关理论与应用, 并且提出了 RBF 网络在控制中的研究与应用新方向.

2　RBF 网络的结构与分类

2. 1　RBF 网络的结构

RBF 神经元网络即 Radial Basis Function Neural Network, 它的产生具有很强的生物学背景. 在人的大脑皮层区域中, 局部调节及交叠的感受野(Receptive Field) 是人脑反应的特点. 基于感受野这一特性, Moody 和 Darken 提出了一种神经网络结构[2, 3], 即 RBF 网络. 图 1 是这种思想的结构图.

这是一种前向网络的拓扑结构, 隐含层的单元是感受野单元, 每个感受野单元输出为

$$\omega = R_i(X) = R_i(\|X - c_i\| / \sigma_i), \quad i = 1, \cdots, H$$

X 是 N 维输入向量, c_i 是与 X 同维数的向量, $R_i(\cdot)$ 具有局部感受的特点. 例如 $R_i(\cdot)$ 取高斯函数, 即 $R_i(X) = \exp(-\|X - c_i\|^2 / \sigma_i^2)$, $R_i(\cdot)$ 只有在 c_i 周围的一部分区域内有较强的反应, 这正体现了大脑皮质层的反应特点. RBF 神经元网络不仅具有上述的生物学背景, 而且还

1996- 10- 10 收稿
[*] 国家自然科学基金资助项目

有数学理论的支持. 文献[4, 5] 利用正则化方法证明了如下结论.

若 $S=\{(X_i, Y_i)\in R^n\times R \mid i=1,\cdots,N\}$ 是训练集合, $\varphi(\cdot, w)$ 表示未知的函数, 其中 w 也未知. 正则化问题的学习过程是寻找 φ及参数 w 使

$$H[\varphi]=\sum_{i=1}^{N}(Y_i-\varphi(X_i, w))^2+\lambda\|P\varphi\|^2$$

最小. 用变分原理可以证明 φ应该选择径向基函数(Radial Basis Function).

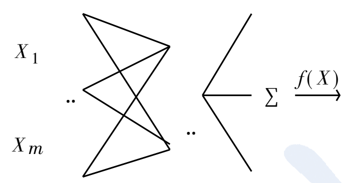

图1　RBF神经网络结构图

2.2　RBF 网络的分类

RBF 网络分类可以从网络参数的不同确定方法入手[6]. RBF 网络的参数确定过程如下[6].

一般RBF 网络可以表示成

$$f_n(X)=\sum_{i=1}^{n}w_i\varphi([X-c_i]^t\Sigma^{-1}[X-c_i])$$

Σ^{-1}是控制 RBF 网络感受野大小的矩阵, 由此可知, 对于给定激活函数 $\varphi(r^2)$ 的 RBF 网络来说有 3 类参数需要确定: (1) RBF 网络权值 w_i, $i=1,\cdots,n$; (2) 网络基函数中心矢量 c_i, $i=1,\cdots n$; (3) 控制矩阵 Σ. 如果用参数 Θ 来代替上述 3 种参数, 针对训练样本 $D_N=\{X_i, Y_i\}_1^N$, 则 RBF 网络的学习过程就是寻找 $\hat{\Theta}$ 使下列式子最小化.

$$\epsilon^2_{\mathrm{RBF}}(D_N, \hat{\Theta})=\min_{\Theta}\epsilon^2_{\mathrm{RBF}}[D_N, f_n(X, \Theta)],\quad D_N=\{X_i, Y_i\}_1^N,$$

$$\epsilon^2_{\mathrm{RBF}}[D_N, f_n(X, \Theta)]=\frac{1}{N}\sum_{i=1}^{N}\left|Y_i-f_n(X, \Theta)\right|^2$$

可以看出对参数 Σ 及 c_i 的优化是很复杂的非线性规划问题, 难以有效的解出. 有效的解决方案是: 首先选定 Σ, 其次从训练样本中确定 c_i, 这样 RBF 网络的参数确定问题就转化为只对输出权值的最小化问题. 这是一个线性方程求解的问题, 因此利用各种最小二乘法都可求出 w_i. 文献[3, 5, 7, 8, 9] 都采用了这种方法.

根据上述参数确定过程将 RBF 网络分类如下:

(1) 0 型 RBF 网络: 如果 RBF 网络的参数确定完全从 $\min_{\Theta}\epsilon^2_{\mathrm{RBF}}[D_N, f_n(X, \Theta)]$ 出发求解 Σ, c_i 及 w_i, 则称这类 RBF 网络为 0 型网络或理想型网络.

(2) Ⅰ型 RBF 网络: 如果 RBF 网络参数中 Σ 预先选定, 中心 c_i 随机从学习样本 X_i 中选取, 则称这类 RBF 网络为Ⅰ型网络或基本型网络.

(3) Ⅱ型 RBF 网络: 如果RBF 网络参数中 Σ 预先选定, 中心 c_i 由聚类方法完成, 则称这类 RBF 网络为Ⅱ型网络或聚类辅助型网络.

3　RBF 网络的函数逼近理论

前向神经网络理论研究的关键问题是函数逼近. 文献[10～17] 在这方面作了深入研究. 文中认为具有 Sigmoid 单元的 3 层前向网络能够逼近连续函数或定义于 R^n 空间紧集上的函数. 这些文献作了一个假设, 就是 Sigmoid 函数必须连续或单调, 而文献[18, 19] 则指出网络逼近函数的能力不是由激活函数的连续性或单调性决定, 而是函数的有界性起了关键作用. 文献

[20] 在 Hornik 工作的基础上[17], 证明了具有局部有界分段连续激活函数的标准前向网络可以逼近任意连续函数至任意精度的充分必要条件是网络的激活函数是非多项式(有限阶次), 并且指出激活函数的阈值有很重要的作用, 它是保证上述充要条件成立的一个重要因素. 文献[21]中, Hornik 指出: (1) 如果激活函数局部黎曼可积且是非多项式, 则前向网络可以一致逼近紧集上的连续函数; (2) 如果激活函数局部有界且是非多项式, 则前向网络可以 $L^p(\mu)$ 逼近连续函数, μ 是具有紧支撑的输入环境测度(Input Environment Measure).

RBF 神经元网络是一类前向网络, 因此上述函数逼近的理论也适用于它. 但针对 RBF 网络有许多特点需要研究. 一般的前向网络(例如多层感知器(MLP))是仿射基函数 ABF (Affine Basis Function) 神经元网络, 这种网络的目标是用函数族 $\sum_{i=1}^{N} w_i g(Y_i X+\theta)$ 逼近连续函数, 而 RBF 网络的目标是用函数族 $\sum_{i=1}^{N} w_i g(\lambda_i \| X- X_i \|_{R^n})$ 逼近连续函数[22]. 文献[22~25]得到了 RBF 网络的一些结果.

定理 1[24] $K: R^n \to R$ 是径向对称、有界可积的函数, 并且满足 $\int_{R^n} K(x)\,\mathrm{d}x \neq 0$ 且几乎处处连续, 那么函数族 $\sum_{i=1}^{N} w_i g\left[\frac{\| x - x_i \|_{R^n}}{\sigma}\right]$ 在 $L^p(R^n)$ 中稠密. 这里 $g(\| x \|_{R^n}) = K(x)$.

定理 2[25] $K: R^n \to R$ 是平方可积函数, 那么函数族 $\sum_{i=1}^{N} w_i g\left[\frac{\| x - x_i \|_{R^n}}{\sigma}\right]$ 在 $L^2(R^n)$ 中稠密的充要条件是 K 是点态的.

Chen T 等人进一步将这个定理一般化为: 假设 $g: R^+ \to R^1$ 满足 $g(\| x \|_{R^n}) \in L^2(R^n)$, 则函数族 $\sum_{i=1}^{N} w_i g\left[\frac{\| x - x_i \|_{R^n}}{\sigma_i}\right]$ 在 $L^2(R^n)$ 中稠密.

文献[22]提出一个连续函数可以作为 RBF 网络激活函数的充分必要条件是函数不能是偶多项式, 并且得到了用 RBF 网络逼近 Banach 空间上的非线性函数及算子的理论结果. 主要有如下 3 个定理.

定理 3 假设 $g \in C(R^1) \cap S'(R^1)$, 且 g 是非偶多项式, X 是 Banach 空间, 它具有 Schauder 基 $\{x_n\}_{n=1}^{\infty}$, $K \subseteq X$ 是 X 中的紧集, f 是定义于 K 上的连续函数. 那么对任意 $\epsilon > 0$, 存在正整数 M, N 及 $y_i^M \in R^M$, 常数 $w_i, \lambda_i \neq 0$, $i = 1, \cdots, N$, 使

$$\left| f(x) - \sum_{i=1}^{N} w_i g(\lambda_i \| x^M - y_i^M \|_{R^M}) \right| < \epsilon$$

对任意 $x \in K$ 成立. 这里 $x^M = [a_1(x), \cdots, a_M(x)] \in R^M$, $x = \sum_{n=1}^{\infty} a_n(x) x_n$.

定理 4 假设 $g \in C(R^1) \cap S'(R^1)$ 是非偶多项式, X 是 Banach 空间, $K \subseteq X$ 是紧集, V 是 $C(K)$ 中的紧集, f 是定义于 V 上的连续函数. 那么对任意 $\epsilon > 0$, 存在正整数 $M, N, x_1, \cdots, x_M \in K$ 及 λ_i, $w_i \in R$, $\xi_i = (\xi_{i1}, \cdots, \xi_{iM}) \in R^M$, $i = 1, \cdots, N$, 使

$$\left| f(u) - \sum_{i=1}^{N} w_i g(\lambda_i \| u^M - \xi_i^M \|_{R^M}) \right| < \epsilon$$

对所有 $u \in V$ 成立. 这里 $u^M = [u(x_1), \cdots, u(x_M)] \in R^M$.

定理 5 假设 $g \in C(R^1) \cap S'(R^1)$ 是非偶多项式, X 是 Banach 空间, $K_1 \subseteq X$, $K_2 \subseteq R^n$ 分别是 X 及 R^n 中的紧集, V 是 $C(K_1)$ 中的紧集, G 是非线性连续算子, 它将 V 映射至 $C(K_2)$ 中, 那

么对任意 $\epsilon > 0$, 存在正整数 M, N, m, 常数

$$w_i^k, \lambda_k, \mu_i \in R,\ k = 1, \cdots, N,\ i = 1, \cdots, M,\ x_1, \cdots, x_m \in K_1,\ \omega_1, \cdots, \omega_N \in R^n$$

使

$$\left| G(u)(y) - \sum_{i=1}^{M} \sum_{j=1}^{N} w_i^k g(\mu_i \| u^m - \xi_{ik}^m \|_{R^m} g(\lambda_k \| y - \omega_k \|_{R^n})) \right| < \epsilon$$

对所有 $u \in V$ 及 $y \in K_2$ 都成立. 这里 $u^m = [u(x_1), \cdots, u(x_m)]$, $\xi_k^n = (\xi_{k1}^m, \cdots, \xi_{km}^m)$, $k = 1, \cdots, N$.

这 3 个定理对于动态系统的逼近有很好的指导作用. 定理 3 主要阐述用频域采样数据逼近动态系统的可能性. 定理 4, 5 阐述了用时域采样数据逼近动态系统的可能性. 上述定理充分说明 RBF 网络对非线性函数进行逼近是切实可行的.

与其他前向网络相比, RBF 网络具有最佳逼近的特性. 在文献[5]中 Poggio 和 Girosi 指出: 从正则理论导出的网络(包括 RBF 网络)具有最佳逼近(Best Approximation)能力. 文中指出采用 BP 算法的多层前向网络不具有最佳逼近性质, 而正则网络(包括 RBF 网络)最佳逼近存在且唯一.

最佳逼近(Best Approximation)定义如下:

定义[26]　函数 $f \in$ 函数集 Φ, $A \subset \Phi$, 距离 $d(f, A) = \inf_{a \in A} \| f - a \|$ 表示 f 与 A 之距离. 如果存在 $a_0 \in A$ 使 $\| f - a_0 \| = d(f, A)$, 则 a_0 是从 A 中对 f 的最佳逼近. 如果对任意 $f \in \Phi$, 至少有一个函数 $a \in A$ 从 A 中最佳逼近 f, 则 A 是存在集. 如果函数 a 是唯一的, 则 A 是唯一集.

对于 BP 型前向网络, 它所表示的函数集合为

$$\sigma^m \equiv \{ f \in C[u] \mid f(X) = \sum_{i=1}^{m} w_i \sigma(XY_i + \theta_i),\ Y_i \in R^d,\ w_i,\ \theta \in R \}$$

理论证明 σ^n 对于 $m \geq 2$ 是非存在集, 因此 BP 型前向网络不存在最佳逼近性质.

对于正则网络, 如 RBF 网, 其表示的函数集合为

$$\sigma^m \equiv \{ f \in C[u] \mid f(X) = \sum_{i=1}^{m} c_i \varphi(X),\ c_i \in R \}$$

其中 $\varphi(X) = G(X, X_i)$, G 是 Green 函数. 对于 RBF 网络, $G = G(\| X - X_i \|)$. 理论证明 T^m 对 $m \geq 1$ 是存在集, 而且如果空间 $C[u]$ 是严格凸的, 则 $T^m (M \geq 1)$ 为存在且唯一集.

由于目前的研究与应用中, 许多神经网络方法都在数字计算机上实现, 文献[27]指出, 如果将数值精度考虑进去, 则会出现一个问题: 最佳逼近的性质不受影响, 而全局逼近(Universal Approximation)的性质将不再存在. 因此最佳逼近的性能体现了 RBF 网络的优点.

由于最佳逼近算法还未成熟, 而全局逼近算法在考虑计算精度时又不能保持良好性能, 那么 RBF 网络该如何解决这个问题呢? 文献[5]说明了 RBF 网络与广义样条函数相当, 它可以用局部逼近的总和达到对训练数据的全局逼近. 因此利用低阶局部逼近的总和可以完成对训练数据的拟合, 这样有限阶次多项式问题就解决了. 这也体现了 RBF 网络比 BP 型网络优越.

RBF 神经网络与 KRE(Kernel Regression Estimator)具有很紧密的联系, 从研究 KRE 入手, 可以得到如下结果[6]:

(1) 得到 RBF 网络函数逼近误差关于隐含单元数目的收敛率的上界;

(2) 给出 RBF 网络一致估计量(Consistent Estimator)的存在性证明;

(3) 当学习样本及隐含单元数目趋于无穷时, RBF 网络最佳相容估计量的逐点收敛率和 L^2 收敛率的上界;

(4)提出了影响 RBF 网络感受野大小恰当选择的因素;

(5)得到了 RBF 网络学习过程中最小二乘估计量经验误差的收敛性结果.

4 RBF 网络的训练方法

RBF 网络学习方法主要有以下几种: (1) Poggio 方法[5]; (2) Moody 和 Darken 方法[3]; (3) 局部学习方法[28]; (4) 正交最小二乘法; (5) 聚类与 Givens 最小二乘联合迭代法.

(1) Poggio 方法是从采用正则方法推导出的. 对训练样本集

$$S = \{(X_i, y_i) \in R^d \times R \mid i = 1, \cdots, n\}$$

只要将集合 S 中 X_i 选为 RBF 网络的中心, 选定 σ, 然后根据 y_i 可以确定出 c_i 的大小, 这样 RBF 网络的训练便可完成.

(2) Moody 和 Darken 的算法与 Poggio 法不同, 其隐含单元数目比训练样本数目少得多. 整个训练过程分为非监督学习和监督学习两个阶段. 非监督学习阶段采用 K-means 聚类方法对训练样本的输入量进行聚类, 找出聚类中心 c_i 及参数 σ, 然后进行监督学习阶段. 由于当 c_i 及 σ_i 确定之后, RBF 网络从输入到输出就成了一个线性方程组, 因此监督学习阶段可以采用最小二乘法求解网络的输出权值 w_i.

(3) 局部学习方法是指 RBF 网络中每个隐含层单元的学习是单独进行的. 假设网络中隐含单元的函数为正交函数, 则从隐层到输出层可以描述为

$$y(X) = \sum_{i=1}^{\infty} w_i g_i(X)$$

$g_i(X) \in L^2[-a, a]$ 为正交函数. 这样网络输出权值 w_i 就可以采用下述方法进行.

(a) $y_i(X) = y(X)$;

(b) 从 $1 \to \infty$ 寻找 j 使 $< y_i(X), g_j(X) >^2$ 最大, 令 $\bar{g}_i(X) = g_j(X)$;

(c) $\begin{cases} w_i = < y(X), \bar{g}_i(X) >; \\ y_{i+1}(X) = y_i(X) - w_i \bar{g}_i(X); \end{cases}$

(d) 若精度不够, $i = i + 1$, 回到(b), 否则终止.

由于 RBF 网络基函数并不满足正交化条件, 因此上述方法需作改进. 局部训练法通过基函数选择来解决这个问题. 基函数选择的原则是: 在空间某个区域内, 只有一个隐含层单元的基函数取值较大, 其他近似为零. 因此这些函数是近似正交的. 为了提高网络的综合性能, 能够适应新的外界输入, 应该加入惩罚项[28]. 这种算法不需要确定函数中心, 也不需要预先知道 RBF 网络隐含单元数目, 因此是一种较好的方法. 但是由于算法中包含复杂的优化问题, 求解有一定难度.

(4) 正交最小二乘法 OLS(Orthogonal Least Squares)是目前训练 RBF 网络应用较多的一种方法. 这种方法的优点是简单易行, 运算速度快, 但是不适合递推运算, 而且基函数中心的确定需要进一步研究. 正交最小二乘法主要由 S. Chen 等人提出, 在文献[7, 29]中分别研究了单输出和多输出的 RBF 网络的正交最小二乘学习方法. 下面是单输出 RBF 网络的 OLS 方法学习过程.

对于具有 N 对学习样本的 RBF 网络, 其输入输出方程为

$$d = p\Theta + E$$

$d = [d(1), \cdots, d(N)^T]$ 为期望输出, $p = [p_1, \cdots, p_N]$ 是隐含单元输出阵, $p_i = [p_i(1), \cdots, p_i$

$(N)^{T}]$, $\Theta=[\theta_1,\cdots,\theta_N]^{T}$为输出权矢量, $E=[\epsilon(1),\cdots,\epsilon(N)]^{T}$是误差矢量.

由此可知 RBF 网络是线性回归模型的一个特例. 回归因子矢量 p_i 构成了基矢量的集合. 由于不同的回归因子一般是相关的, 因此需要有一个方法来确定各个不同的回归因子对输出能量的贡献, OLS 方法解决了这个问题. OLS 方法的过程是将 p_i 集合变换成正交集, 即 $p=WA$, A 是对角线为 1 的上三角阵, W 是包含正交矢量 w_i 的矩阵. 因此有方程 $d=Wg+E$, $A\Theta=g$. 由于 A, g 可以在正交化的过程中得到, 因此 Θ 也易求解出来. OLS 法在每步正交过程中, 都要用"新息-贡献"准则进行正交优选. 当 OLS 法满足一定精度后, 算法即终止, 这时的正交矢量数目就是隐含层神经元数目.

(5) Givens 迭代算法是由 S. Chen 等人提出. 文献[30]介绍了 Givens 迭代算法, 并且将其与聚类方法联合得到了一种较为有效的 RBF 网络训练算法. 但这种方法运算的存储量大, 运算速度慢, 对实时运行有一定影响. Givens 迭代算法包含聚类和 Givens 最小二乘法有如下目的: (a) 对于 RBF 网络的实时应用, 由于输入数据不断变化, 因此 RBF 网络的函数中心需要恰当选择以使它对输入域充分采样, 并且可以跟踪上输入数据的变化模式, 所以算法采用了递推 K-means 聚类方法; (b) 利用 Givens 最小二乘法的良好性能使 RBF 网络的权值得到实时修正.

5　RBF 神经网络的优点及问题

RBF 网络是一种性能良好的前向网络. 它不仅有全局逼近性质, 而且具有最佳逼近性能. RBF 网络结构上具有输出-权值线性关系, 同时训练方法快速易行, 不存在局部最优问题[31]. 这些优点给 RBF 网络的应用奠定了良好的基础.

尽管如此, RBF 网络有许多问题需要解决.

(1) 如何确定网络激活函数的数据中心. 目前许多方法都从聚类出发, 但是聚类需要有一个度量问题. 如何定义这种度量才能恰当的找到 RBF 网络的数据中心还需研究, 因为在仿真研究中发现 RBF 网络的数据中心对 RBF 网络的学习速度及性能有较大影响.

(2) 如何寻找合适的径向基函数. 对于一组给定的学习数据, 往往反映了很复杂的非线性关系, 而且数据相关性较大, 如果基函数选择不当, 那无论怎么改进 RBF 网络的学习方法, 都难以达到学习精度或根本不能完成学习任务.

(3) 研究 RBF 网络最佳逼近算法.

(4) 研究 0 型 RBF 网络训练方法. 因为目前 RBF 网络的主要训练方法都是针对 I 型和 II 型 RBF 网络的, 往往实际效果不如理论所述那么理想, 主要原因是 I 型和 II 型 RBF 网络没有考虑在网络训练中寻找最优的函数中心 c_i 和控制矩阵 Σ.

(5) 设计快速有效的迭代算法训练 RBF 网络. 迭代算法是网络实时运行的需要, 而目前已有的迭代方法存储量大, 运算较慢, 所以需要快速的迭代算法. 最近我们正在研究用 PLS[32] 方法来训练 RBF 网络, 以期达到网络实时运行的效果.

6　RBF 神经网络在控制中的应用

RBF 网络存在许多优点, 在不少领域中得到了应用. 目前在控制领域中 RBF 网络主要在以下两方面应用较多: (1) 系统辨识与建模; (2) 控制方案设计.

6.1　系统辨识与建模

前向网络应用于系统辨识和建模有一个标准模式, 即基于系统输入输出数据完成系统辨识, 其框图如图 2 所示.

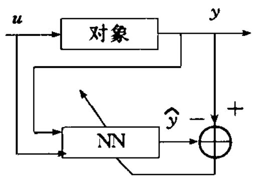

图 2　前向网络辨识系统的框图

因此 $\hat{y}(k+1) = \mathrm{NN}[y(k), y(k-1), \cdots, y(k-l+1), u(k), u(k-1), \cdots, u(k-m+1)]$ 完成了对系统的辨识.

$$\begin{cases} x(k+1) = f(x(k), u(k)) \\ y(k) = h(x(k)) \end{cases}$$

完成这样一个系统的辨识通常作以下假设[33]: (1) f, h 是光滑的; (2) 系统是状态可逆的; (3) 系统阶次的上界可知. 文献[33] 对这些问题作了研究, 并且指出全局的输入输出模型对一般可观性(Generic Observability) 系统都存在, 而并不要求系统的强可观性(Strong Observability). 这说明 RBF 网络可以对几乎所有的系统进行辨识和建模.

RBF 网络用于非线性系统辨识主要优势在于: 它可以避开复杂的算法而较为准确的完成辨识任务.

文献[34] 对随机动态系统的逼近和状态估计作了研究. 文献[35] 用 RBF 网络对非线性系统进行了建模仿真研究, 文中讨论了 RBF 网络选用 RM (Reciprocal Multiquadric) 函数作为激活函数的优点, 并且将其应用于连续搅拌反应釜和 pH 中和过程的建模研究, 同时还研究了模型稳定性问题.

RBF 网络用于非线性系统辨识和系统建模一般分为以下几个步骤.

(1) 恰当选择学习样本. 在许多文献中, 系统辨识的学习数据都用伪随机码激励系统得到, 在过程控制中, 这是不适用的. 无论采用什么方法得到学习数据都必须遵循一条原则, 即学习样本必须充分体现系统的工作状况.

(2) 学习样本数据的处理. 一般来说学习数据都应做归一化处理, 同时由于在实时控制中采集到的数据含有噪声, 因此需要有滤波的处理过程.

(3) 确定模型的阶次. 这可以应用被建模系统的先验知识来确定, 也可通过数据分析得到.

(4) 采用恰当的学习算法完成 RBF 网络的离线学习.

(5) 如果系统是时变的, 必须用递推算法对 RBF 网络进行在线校正.

6.2　控制方案设计

神经网络应用于控制方案设计非常普遍, 而 RBF 网络应用于控制方案设计还不常见. 文献[36] 提出基于 RBF 网络的自适应控制, 文献[37] 提出了基于 RBF 网络的预测控制方案, 文献[38] 采用两个不同的高斯网络(RBF 的一种) 提出了非线性内模控制, 文献[39] 指出这种内模控制是有余差控制, 因此 Seborg 等人在文献[40] 中利用 RBF 网络设计了一种无余差的内模控制, 并且将其应用于 pH 中和过程的控制. 文献[41] 中用 RBF 网络作为预测控制的预测模型, 用 Hopfield 网络作为控制器, 从理论上设计了一种非线性预测控制方案, 但是这种方案还缺乏仿真结果. 文献[42] 用 RBF 网络设计了非线性随机系统的最优控制.

图 3 是预测控制[37] 和无余差内模控制[40] 的模式.

6.2.1　基于 RBF 网络的非线性预测控制方案

预测控制有 3 要素[43]: 预测模型, 滚动优化, 反馈校正. 在 RBF 网络设计预测控制器时, 可以借助于 RBF 网络的优点完成这些任务. 特别是非线性系统预测控制, 预测模型一般难以得到, 同时在每个采样间隔内完成的计算量特别大, 这不利于预测控制的实际运用.

文献[37]用 RBF 网络作预测模型, 而且预测控制器也用 RBF 网络实现, 该文对这种控制方案作了仿真研究, 但没有给出整个控制系统的稳定性和算法收敛性说明.

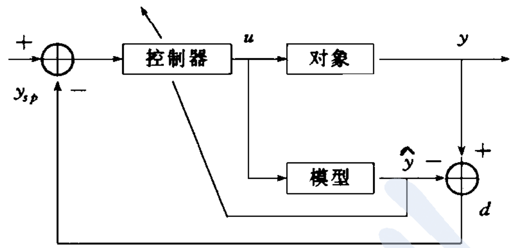

图 3　预测控制与内模控制模式

6.2.2　基于 RBF 网络的无余差内模控制

由于许多过程可以用这样的模型表示: 模型的最当前输入与模型输出是线性关系. 因此利用两个RBF 网络可以表示这类过程.

$$\hat{y}(k) = F_1(\hat{y}(k-1),\cdots,\hat{y}(k-m),u(k-\theta-2),\cdots,u(k-\theta-n)] + u(k-\theta-1) + F_2(\hat{y}(k-1),\cdots,\hat{y}(k-m),u(k-\theta-2),\cdots,u(k-\theta-n)]$$

$F_1[\cdot] = a_{01} + \sum_{i=1}^{M_1} a_{i1}\varphi(\|\hat{x}(k) - c_{i1}\|)$, $F_2[\cdot] = a_{02} + \sum_{i=1}^{M_2} a_{i2}\varphi(\|\hat{x}(k) - c_{i1}\|)$是两个 RBF 网络.

$$\hat{x}(k) = [(\hat{y}(k-1),\cdots,\hat{y}(k-m),u(k-\theta-2),\cdots,u(k-\theta-n)]^T$$

根据上述模型设计控制器如下.

首先确定控制指标为

$$I = [(\hat{y}(k+\theta+2) - s(k)]^2 + \gamma[u(k) - u(k-1)]^2$$

$$s(k) = \alpha\hat{y}(k+\theta) + (1-\alpha)[r(k) - d(k)],\ 0 < \alpha < 1$$

式中$r(k)$是设定值, $d(k) = y(k) - \hat{y}(k)$.

其次设计控制器, 使上述指标极小化.

$$u(k) = \frac{s(k)F_2 - F_1F_2 + \gamma u(k-1)}{F_2^2 + \gamma}$$

虽然 F_1, F_2, s 都包含模型的将来输出, 但是控制律还是可行的, 因为这些将来输出都可以用 RBF 网络模型预测出来. 可以证明上述控制是无余差控制.

7　RBF 神经网络在控制中的研究及应用新方向

7.1　控制方案设计

应用神经网络设计控制器, 主要有两种方向: (1) 先辨识对象模型, 然后针对对象模型设计控制器, 控制器可以是神经网络, 也可以是常规控制器; (2) 直接由神经网络设计控制器, 这是一种由网络学习系统逆动力学行为的方法. 由于一个实际的物理系统, 逆动力学行为并不一定存在, 因此较多方法从第一种思路入手, 设计神经网络控制方案.

在作者研究中[41], 从理论上推导出了基于 RBF 网络预测模型的非线性预测控制方案. 由于用 RBF 网络作为预测模型, 在线滚动优化难以进行, 因此采用 Hopfield 网络完成这个优化问题. 但是这种方案计算量大, 因此一直没有成功地仿真出来. 作者注意到如果采用线性化的设计方法可以解决这个问题. 文献[44]采用模型线性化的方法完成了神经网络自校正控制器设计.

神经网络控制器设计遇到的关键问题是很难对系统的稳定性和收敛性进行分析. 不过文献[45, 46]在这方面作了一定研究, 但还是非常有限, 今后这方面的工作需要进一步深入.

7.2 RBF 网络对非线性动态系统的建模

非线性系统建模如果采用常规算法是一个非常棘手的问题, RBF 网络在这个方面很有优势, 只要给出被建模对象充分完备的数据, RBF 网络能够很容易将系统模型建立出来.

一般RBF 网络应用于动态系统建模, 需要根据 ARMA 模型的形式将静态的 RBF 网络转化为具有动态特性的网络. 如图 4 和图 5 两种方式.

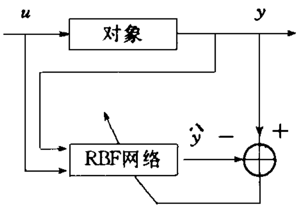

图 4 RBF 网络用于动态系统建模(方案 1)

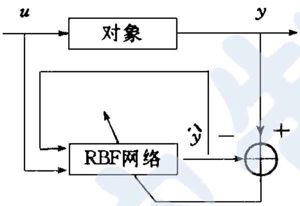

图 5 RBF 网络用于动态系统建模(方案 2)

用 RBF 网络对非线性系统建模需要解决的关键问题是样本数据的选择. 在仿真研究中我们可以通过激励系统而得到系统非常完备的信息, 这样对象的模型就可以较为准确地建立起来. 作者采用这种方法对化工中的连续搅拌反应釜作了建模研究, 并且利用阶跃响应对模型进行校验, 发现建模效果相当好. 可是遗憾的是在实际工业过程中, 系统的信息往往只能从系统运行的操作数据中分析得到. 因此如何从系统运行的操作数据中提取系统运行状况信息, 需要进一步研究.

由于实际系统往往是时变的, 因此设计有效的在线 RBF 网络建模方法也是今后研究的方向.

7.3 RBF 神经网络在软测量中的应用

软测量方法是工业过程控制的重要工具. 由于许多工业过程的重要输出变量很难通过传感器得到, 如果采用在线分析仪则需较大投资, 并且测量滞后非常大, 不利于过程的在线监视和控制. 针对这些问题, 文献提出了推断控制方案, 其基本思想是: 根据比较容易测量的过程辅助变量(2 次测量变量) 来估计并克服不可测扰动对过程主要输出变量的影响. 推断控制设计包括估计器和控制器设计. 估计器设计就是软测量方法. 文献[48] 指出软测量是过程控制中主要方向.

软测量涉及到建模、辨识和数据处理问题, 由于工业过程中许多系统是非线性的, 因此我们采用 RBF 网络来实现过程的软测量.

软测量模型建立需要注意以下问题[49]: (1) 2 次测量变量的选择; (2) 输入数据的处理. 这两方面问题解决之后才能有效的建立软测量模型. 在理论建模研究中, 这容易忽视, 而在实际过程控制中, 这部分工作相当重要, 如果这部分工作做得好, 那么就可以得到很好的建模数据, 软测量模型就较易建立.

采用 RBF 网络建立软测量模型是过程非线性和复杂性的需要. 至今我们已经用 RBF 网络建立了化工中精馏塔系统的软测量模型.

软测量模型需要不断维护, 这要求软测量模型能够在线校正, 因此研究 RBF 网络的在线训练方法对软测量具有很重要的作用.

7.4 RBF 神经网络与模糊技术

文献[50]证明了RBF 网络与模糊推理系统(Fuzzy Inference System)具有函数等价性,这启发我们可以从以下两方面作一些有意义的工作.

(1)利用模糊推理系统确定RBF 网络的结构及参数.文献[44]利用模糊技术构造了一种前向网络,并且将其应用于控制器设计.这充分说明用模糊技术训练RBF 网络的可行性.在以后的研究中,我们将运用模糊技术实现RBF 网络的学习过程.如果能够成功,那将会大大促进RBF 网络的发展.

(2)利用RBF 神经网络建立模糊推理系统.模糊推理系统由模糊规则集、包含语义表隶属度函数的知识库及推理机组成.文献[51～53]提出了将模糊推理系统转换成函数等效的自适应神经网络,然后利用BP 算法来更新前提参数(Premise Parameters),用最小二乘法来更新结果参数(Consequent Parameters).文献[54]提出了用神经网络建立模糊知识库和专家系统.可见神经网络建立模糊推理系统是有效的.由于RBF 网络比一般前向网络更优越,因此可以研究用RBF 网络建立模糊系统,最终的目的是实现系统的模糊控制.

7.5 基于记忆的RBF 网络在分类器中的应用及其对操作优化的作用

文献[55]提出一种好的神经网络理论必须遵循如下准则:

(1)完成网络的设计任务(Perform Network Design Task);

(2)学习鲁棒性(Robustness in Learning);

(3)学习快速性(Quickness in Learning);

(4)学习有效性(Efficiency in Learning);

(5)学习推广性(Generalization in Learning).

从这些原则出发,文中提出了一种产生RBF 网络结构同时完成网络学习的新方法.这种方法主要应用于解决分类问题.

由于许多神经元网络学习算法是无记忆的,这会降低网络的学习效率,尤其在解决分类问题时,更是如此.与无记忆神经网络相比,人脑学习是有记忆的.对新输入的数据,人脑往往先与记忆中的知识作比较,然后重新学习.这种思想非常有利于实时学习,因此文献[55]提出了一种在线自适应的学习方法,这种方法力图使RBF 网络具有记忆,然后根据这种思想提出一种实时分类方法.

分析这种有记忆的神经网络学习思想主要目的在于:试图将基于记忆的RBF 网络应用于系统的操作优化中去.文献[56]利用模式识别的方法对乙烯生产过程的操作优化问题进行了求解,利用操作点的优化提高乙烯产品质量.纵观全文,我们完全可以用分类方法去解决同一问题,而且利用基于记忆的RBF 网络可以完成分类的在线运行,这样操作优化就可以实时进行.这不但是对文献[56]的很大改进,而且对实际应用非常有价值,因为目前许多操作优化方法还只是离线进行.目前这种在线优化思想正在完善之中,估计不久将得到仿真结果.

参 考 文 献

1 Antsaklis P J. Neural Networks in Control Systems, Special Section on Neural Networks for Systems and Control. IEEE Control System Magazine, 1990: 3～5

2 Moody J, Darken C. Learning with Localized Receptive Fields. In Proc 1988 Connectionist Models Summer School, D Touretzky, G Hinton, and T Sejnowski (Eds.), Carnegie Mellon University, Morgan Kaufmann Publishers, 1988

3 Moody J, Darken C. Fast Learning in Networks of Locally-tuned Processing Units. Neural Computation, 1989, (1):

281～294

4 Kung S Y. Digital Neural Networks. PTR Prentice-Hall Inc, 1993: 175～179

5 Piggio T, Girosi F. A Theory of Networks for Approximation and Learning. A I Memo No 1140, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge Mass, 1989

6 Xu Lei, *et al*. On Radial Basis Function Nets and Kernel Regression: Statistical Consistency, Convergence Rates and Receptive Field Size. Neural Networks, 1994, **7**(4): 609～628

7 Chen S, Grant P M, Cowan C F N. Orthogonal Least Squares Algorithm for Training Multioutput Radial Basis Function Networks. IEE Proceedings-F, 1992, **139**(6): 378～384

8 Broomhead D S, Lowe D. Multivariable Functional Interpolation and Adaptive Networks. Complex Systems, 1988, (2): 321～323

9 Powell M J D. Radial Basis Functions for Multivariable Interpolation: A Review. In J C Mason & M G Cox(Eds.) Algorithms for Approximation, Oxford: Clarendon Press

10 Wieland A, Leighten R. Geometric Analysis of Neural Network Capacity. In Proc IEEE 1st ICN, 1987, (1): 385～392

11 Irie B, Miyake S. Capacity of Three-layered Perceptrons. In Proc IEEE ICNN, 1988, (1): 641～648

12 Carroll S M, Dtckinsong B W. Construction of Neural Nets Using Random Transform. In Proc ICNN, 1989, (1): 183～192

13 Funahashi K. On the Approximate Realization of Continuous Mappings by Neural Networks. Neural Networks, 1989, (2): 183～192

14 Mhaskar H N, Micchelli C A. Approximation by Superposition of Sigmoidal and Radial Functions. Advances on Applied Mathematics, 1992, **13**: 350～373

15 Cybenko G. Approximation by Superpositions of a Sigmoidal Function. Mathematics of Control, Signals and Systems, 1989, **2**(4): 303～314

16 Ito Y. Representation of Functions by Superpositions of a Step of Sigmoidal Function and Their Applications to Neural Network Theory. Neural Networks, 1991, **4**: 385～194

17 Hornik K. Approximation Capabilities of Multilayer Feedforward Networks. Neural Networks, 1991, **4**: 251～257

18 Chen T, Chen H, Liu R W. Approximation Capacity in C(R^n) by Multilayer Feedforward Networks and Related Problems. IEEE Trans on Neural Networks, 1995, **6**(1)

19 Chen T, Chen H, Liu R W. A Constructive Proof of Cybenko's Approximation Theorem and Its Extensions, Computing Science and Statistics. Lepuge and Page (Eds.) In Proc 22nd Symp Interface, East Lansing, Michigan, 1990: 163～168

20 Moshe Leshno, *et al*. Multilayer Feedforward Networks with a Nonpolynomial Activation Function Can Approximate Any Function. Neural Networks, 1993, **6**: 861～867

21 Hornik K. Some New Results on Neural Network Approximation. Neural Networks, 1993, **6**: 1069～1072

22 Chen Tian Ping, Chen Hong. Approximation Theory Capability to Functions of Several Variables, Nonlinear Functionals, and Operators by Radial Basis Functional Neural Networks. IEEE Trans on Neural Networks, 1995, **6**(4): 904～910

23 Lippmann R P. Pattern Classification Using Neural Networks. IEEE Commun Mag, 1989, **27**: 47～64

24 J Park, Sandberg I W. Universal Approximation Using Radial Basis Function Networks. Neural Computation, 1991, **3**: 246～257

25 Park J, Sandberg I W. Approximation and Radial Basis Function Networks. Neural Computation, 1993, **5**

26 Girosi F, Poggio T. Networks and the Best Approximation Property. Biological Cybernetics, 1990, **63**: 169～176

27 Jonathan Wray, Gray G R Green. Neural Networks, Approximation Theory and Finite Precision Computation. Neural Networks, 1995, **8**(1): 31～37

28 Tyler Holcomb, Manfred Morari. Local Training for Radial Basis Function Networds: Towards Solving the Hidden Unit Problem. American Control Conference, 1991: 2331～2335

29 Chen S, Cowan C F N, Grant P M. Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks.

IEEE Trans on Neural Networks, 1991, **2**(2): 302～309

30 Chen S, Billings S A, Grant P M. Recursive Hybrid Algorithm for Nonlinear System Identification Using Radial Basis Function Networks. Int J Control, 1992, **55**(5): 1051～1070

31 Monica Bianchini, *et al*. Learning without Local Minima in Radial Basis Function Networks. IEEE Trans on Neural Networks, 1995, **6**(3): 749～755

32 Qin S J, McAvoy T J. Nonlinear PLS Modeling Using Neural Networks. Computer & Chemical Engineering, **16**(4): 379

33 Asriel U Levin, Kumpati S Narendra. Recursive Identification Using Feedforward Neural Networks. Int J Control, 1995, **61**(3): 533～547

34 Suni Elanayar V T, Yung C Shin. Radial Basis Function Neural Network for Approximation and Estimation of Nonlinear Stochastic Dynamic Systems. IEEE Trans on Neural Networks, 1994, **5**(4): 584～603

35 Martin Pottmann, Dale Seborg. Identification of Nonlinear Process Using Reciprocal Multiquadric Functions. J Proc Cont, 1992, **2**(4): 189～203

36 Sanner R M, *et al*. Gaussian Networks for Direct Adaptive Control. IEEE Trans on Neural Networks, 1992, **3**: 837～863

37 Pottmann M, Seborg D E. A Nonlinear Predictive Strategy Using Radial Basis Function Networks. Preprints of the IFAC Symposium DYCORD, College Park, MD, 1992: 309～314

38 Hunt K J, Sburburo D. Neural Networks for Nonlinear Internal Model Control. IEE Proc-D, 1991, **138**: 431～438

39 Nahas E, Henson M A, Seborg D E. A Nonlinear Internal Model Control Strategy for Neural Network Models. Comp Chem Eng, 1992, **16**: 1059～1057

40 Martin Pottmanm, Dale E Seborg. A Radial Basis Function Control Strategy and Its Application to a pH Neutralization Process. The Second European Control Conference ECC'93, Groningen, the Netherlands, 1993

41 王旭东. 用径向基函数神经元网络进行非线性系统的建模与控制. 上海交大硕士论文, 1995

42 Parisini T, Zoppoli R. Radial Basis Functions and Multilayer Feedforward Neural Networks for Optimal Control of Nonlinear Stochastic Systems. International Joint Conference on Neural Networks, 1993: 1853～1858

43 席裕庚. 预测控制. 国防工业出版社, 1993

44 王殿辉, 柴天佑. 自适应模糊神经网络控制器设计的线性化方法. 控制与决策, 1995, **12**(1)

45 任雪梅, 高为炳. 基于神经网络非线性与系统辨识和控制的研究. 控制理论及应用, 1995, **12**(2): 147～153

46 任雪梅, 高为炳. 一类非线性系统神经网络控制的稳定性分析. 控制与决策, 1995, **10**(2): 465～469

47 Joseph B C, Brosillow C B. Inference Control of Process. AIChEJ, 1978, **24**(3)

48 McAvoy T J. Contemplative Stance for Chemical Process. Automatica, 1992, **28**(2): 441～442

49 罗荣富, 邵惠鹤. 软测量方法及其工业应用. 第六届过程控制科学报告会, 苏州, 1993: 324～329

50 Roger J S, Sun C T. Functional Equivalence between Radial Basis Function Networks and Fuzzy Inference Systems. IEEE Trans on Neural Networks, 1995, **4**(1): 156～158

51 Roger J S. Fuzzy Modeling Using Generalized Neural Networks and Kalman Filter Algorithm. In Proc Ninth National Conf Artificial Intelligence, 1991: 762～767

52 Roger J S. Rule Extraction Using Generalized Neural Networks. In Proc 4th IFSA World Congress, 1991

53 Roger J S. ANFIS: Adaptive-network-based Fuzzy Inference Systems. IEEE Trans on System, Man, Cybern, 1991

54 Ronald R Yager. Modeling and Formulating Fuzzy Knowledge Basis Using Neural Networks. Neural Networks, 1994, **7**(8): 1273～1283

55 Asim Roy, *et al*. An Algorithm to Generate Radial Basis Function (RBF)-like Nets for Classification Problems. Neural Networks, 1995, **8**(2): 179～201

56 Huang Sunan, Shao Huihe. Application of Pattern Recognition to Ethylene Production Optimization. Eng Applic Artif Intell, 1994, **7**(3): 329～333

THE THEORY OF RBF NEURAL NETWORK AND ITS APPLICATION IN CONTROL

WANG Xudong SHAO Huihe
(*Institute of Automation, Shanghai Jiao Tong University* 200030)

Abstract This paper first surveys the structure, types, approximation theory, and training methods of the RBF neural network, then analyzes the advantages and problems. At the same time, some typical applications in control have been discussed. At last the new trends of the applications in control are pointed out.

Key words RBF neural networks, universal approximation, best approximation, predictive control, internal model control, soft sensor, optimization of operation

作者简介

王旭东, 男, 26 岁, 博士研究生. 研究领域为过程控制, 智能控制等.

邵惠鹤, 男, 61 岁, 教授, 博士生导师. 研究领域为过程控制, 生化控制, 智能控制等.