

尺度可调的混合核 RBF 网络

付丽华¹, 李宏伟¹, 张 猛²

(1. 中国地质大学数学与物理学院, 湖北武汉 430074; 2. 华中师范大学计算机科学系, 湖北武汉 430079)

摘 要: 针对传统核模型中采用单一核函数的局限性, 利用两个核函数的线性组合得到混合核. 在 RBF 网络的训练中, 采取正交最小二乘的方法进行逐步回归建模. 在学习每个神经元参数时, 首先, 用全局 k 均值聚类法得到数据样本的聚类中心, 然后对每一个聚类中心, 利用群搜索优化器搜索出最佳的尺度和混合核调节参数, 误差最小的参数组合即为径向基函数参数. 实验说明, 新的 RBF 网络具有稀疏性好, 泛化能力高等优点.

关键词: 混合核; RBF 网络; 群搜索优化器; 正交最小二乘

中图分类号: TP301.6 **文献标识码:** A **文章编号:** 0372-2112 (2011) 01-0184-06

Mixture Kernel RBF Network with Tunable Scales

FU Li-hua¹, LI Hong-wei¹, ZHANG Meng²

(1. School of Mathematics and Physics, China University of Geosciences, Wuhan, Hubei 430074, China;

2. Department of Computer, Central China Normal University, Wuhan, Hubei 430079, China)

Abstract: To overcome the limitation of one single kernel in the traditional kernel function model, a new type of mixture kernel is constructed by combining two kernel functions linearly. Each individual regressor in RBF network is trained term by term using orthogonal least squares algorithm. In the training phase, a global k -means cluster algorithm is used to decide the kernel centers. For each clustering centers, Group search optimizer (GSO) is utilized to get the proper scale and weight within the kernel. The optimal kernel parameters which minimize the training error are used to shape each regressor. The experiments results show that the new RBF network is sparser than some traditional RBF network with one single kernel.

Key words: mixture kernel; radial basis function (RBF) network; group search optimizer; orthogonal least squares

1 引言

RBF 神经网络(radial basis function neural network)是一种性能良好的前向网络, 具有最佳逼近及克服局部极小值问题的性能, 在函数逼近、模式识别、动态系统控制^[1]等方面都有广泛的应用. 但是与 BP 网络相比, 它需要更多的隐层神经元, 使网络在训练样本时易产生过拟和. 因此, 应尽量减小 RBF 网络规模, 增强模型的稀疏性, 以提高网络泛化能力. 稀疏性通常可以用 l_p ($p \leq 1$) 模或信号分量(分解系数)按绝对值排序后向零值衰减的速度来度量^[2].

对于 RBF 神经网络的学习算法, 关键是三组参数的确定: 核函数中心、尺度及对应的权值. 这些参数或者分开确定, 或者同时确定. 通常, 在分开确定参数时, 往往先决定中心和尺度, 然后估计权值. 当中心和尺度候选项的集合确定时, 正交最小二乘法(Orthogonal Least Squares, OLS)可以有效的构造出较为稀疏的网络^[3]. 此外, 一些类似于子空间选择方法也可以用于构造网络.

同时确定三组参数时, 可以用有指导的学习策略, 根据非线性优化准则进行参数的估计. 例如, 1999 年, Karayiannis^[4]提出的基于梯度下降搜索法 (gradient descent search) 的 RBF 神经网络构造方法; 2003 年, Lazaro 等^[5]提出的期望最大估计 (expectation-maximization) 法和 2003 年 Gonzalez 等人提出的演化算法^[6]构造网络.

在已有的常用学习算法中, 核函数中心参数(或者中心参数的备选项)要么是从给定的训练样本集里按照某种方法直接选取, 要么采用聚类的方法进行确定. 然而, 尺度的确定却没有简单有效的方法. 传统的单尺度 RBF 网络中, 所有的核函数用同一个固定的尺度, 这种网络的特点是构造方法简单, 但结果是模型非稀疏, 泛化能力很差. Billings 等人于 2007 年提出一种多尺度 RBF 网络(MSRBF)^[7], 经验的给出一个多尺度的候选项集合, 然后运用正交最小二乘法逐项从中选择“最优”的尺度. 这种方法的优点是计算简单, 但是由于尺度候选项集合是事先经验构造, 仍然存在一定的局限性.

RBF 网络一般采用单一核函数, 其中最常用的核函

数是高斯核函数.但是当数据集合的分布较为复杂时,单个核函数往往不能完全涵盖数据的分布特性,此时单核 RBF 网络模型需要更多的神经元来表示数据特征,这将破坏了模型的稀疏性.这个问题也引起了统计机器学习领域的重视,Ying Tan 等人提出了混合核支持向量机(Support Vector Machines with Hybrid Kernel, SVMHK)^[8].与经典 SVM 不同,SVMHK 采用了多种不同核函数的线性组合作为新的核函数.实验表明,核函数的组合能给模型带来较好的泛化能力,提高模型的性能.

本文提出一种可调尺度的混合核 RBF 神经网络,将多个核函数线性组合成一个混合核,通过参数调节各个核函数的权重;并利用全局 k 均值聚类算法计算出样本的聚类中心,作为核函数中心参数的候选集合;混合核调节参数和尺度则通过演化算法领域的最新研究成果群搜索优化器(Group Search Optimizer, GSO)^[9]搜索得到.GSO 受到鱼群觅食启发,包括复制器(Producer)选择(离食物最近的鱼选择,即最优鱼选择);复制器复制(最优鱼向前,向左或向右觅食),搜索器(Scrounger)选择与搜索(部分鱼群向最优鱼方向靠拢);分散(Dispersed)(部分鱼群随机游动).其基本思想如图 1 所示.GSO 是一种针对连续函数的优化算法且特别适合于高维,多原层(multimodal)的优化问题.关于 GSO 的具体算法可参考文献[9].最后,利用正交最小二乘法完成网络的训练.

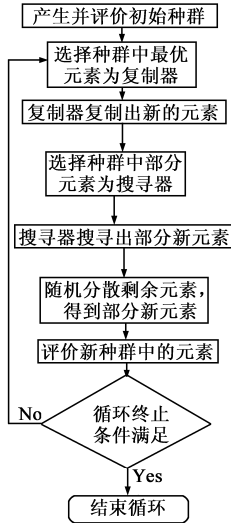


图1 GSD算法的流程

器(Scrounger)选择与搜索(部分鱼群向最优鱼方向靠拢);分散(Dispersed)(部分鱼群随机游动).其基本思想如图 1 所示.GSO 是一种针对连续函数的优化算法且特别适合于高维,多原层(multimodal)的优化问题.关于 GSO 的具体算法可参考文献[9].最后,利用正交最小二乘法完成网络的训练.

2 回归模型及正交最小二乘法

考虑如下非线性的离散时间系统 NARX 模型^[7]:

$$y(t) = f(y(t-1), \dots, y(t-n_y), u(t-1), \dots, u(t-n_u)) + e(t) \quad (1)$$

其中, $u(t)$, $y(t)$, $e(t)$ 分别是系统的输入、输出和噪声干扰项; n_u , n_y 分别是输入 $u(t)$ 和输出 $y(t)$ 的最大延迟, 为正整数; $f(\cdot)$ 是未知的非线性映射; 噪声 $e(t)$ 是白噪声序列.

令 $d = n_u + n_y$, $\mathbf{x}(t) = [x_1(t), \dots, x_d(t)]^T$, 其中

$$x_k(t) = \begin{cases} y(t-k), & 1 \leq k \leq n_y \\ u(t-(k-n_y)), & n_y+1 \leq k \leq n_y+n_u \end{cases}$$

这样,模型(1)可以看成由数据对 $\{\mathbf{x}(t), y(t)\}_{t=1}^N$ 经过某种非线性映射 $f(\cdot)$ 得到.

一种常用的回归模型:

$$y(t) = \hat{f}(\mathbf{x}(t)) + e(t) = \sum_{m=1}^M w_m \phi_m(\mathbf{x}(t)) + e(t)$$

其中, M 是回归子的个数, $\phi_m(\mathbf{x}(t))$ 为模型核函数, w_m 是权系数 ($m=1, 2, \dots, M$). 令输出向量 $\mathbf{y} = [y(1), \dots, y(N)]^T$, 回归模型的矩阵形式如下

$$\mathbf{y} = \Phi \mathbf{w} + \mathbf{e} \quad (2)$$

其中 $\mathbf{w} = [w_1, \dots, w_M]^T$, $\mathbf{e} = [e(1), \dots, e(N)]^T$ 为误差向量, 表示每个数据点的误差. $\Phi = [\phi_1(\mathbf{x}(t)), \dots, \phi_M(\mathbf{x}(t))]^T$. 目标是选择最优的 $\{\phi_i | i=1, \dots, M\}$ 的线性组合, 根据某种准则来近似估计输出向量 \mathbf{y} .

正交最小二乘法(OLS)在变换的正交空间中求解, 即将回归矩阵 Φ 进行正交分解,

$$\Phi = \mathbf{H} \mathbf{A} \quad (3)$$

\mathbf{A} 是一个单位对角的上三角矩阵, 而 $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_M]$ 是一个任意两列正交的矩阵, 即 $\mathbf{H}_i^T \mathbf{H}_j = 0$, 如果 $i \neq j$. 这样回归模型(2)变为

$$\mathbf{y} = \mathbf{H} \boldsymbol{\theta} + \mathbf{e} \quad (4)$$

$\boldsymbol{\theta} = \mathbf{A} \mathbf{w}$ 表示新的权向量, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_M]^T$.

容易证明, 回归模型均方误差 (Mean Square Error, MSE) 可以表示为

$$J = \mathbf{e}^T \mathbf{e} / N = \mathbf{y}^T \mathbf{y} / N - \sum_{i=1}^M \mathbf{H}_i^T \mathbf{H}_i \theta_i^2 / N \quad (5)$$

利用贪婪算法, 正交最小二乘法利用最小化式(5)的方法逐项寻找最优解 \mathbf{H}_i , θ_i . 第 L 步的回归模型可以表示为

$$\hat{\mathbf{y}} = \hat{\mathbf{H}} \hat{\boldsymbol{\theta}} \quad (6)$$

其中 $\hat{\mathbf{H}} = [\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_L]$, $\hat{\boldsymbol{\theta}} = [\theta_1, \dots, \theta_L]^T$.

3 可调尺度的混合核 RBF 网络

3.1 混合核

将多个不同的核函数结合起来, 使得混合后的核函数具有更适合非线性系统的建模问题, 这是混合核方法的基本思想.

RBF 网络中最常用的是高斯核函数, d 维高斯核函数的形式为:

$$\phi_i^{(G)}(\mathbf{x}(t); \boldsymbol{\sigma}_i, \mathbf{c}_i) = \exp\left\{-\frac{1}{2} [\mathbf{x}(t) - \mathbf{c}_i]^T \boldsymbol{\Lambda}_i^{-1} [\mathbf{x}(t) - \mathbf{c}_i]\right\}$$

其中, $\mathbf{x}(t) = [x_1(t), \dots, x_d(t)]^T$ 是一个 d 维输入向量, $\mathbf{c}_i = [c_{i,1}, \dots, c_{i,d}]^T$, $\boldsymbol{\sigma}_i = [\sigma_{i,1}, \dots, \sigma_{i,d}]^T$ 和 $\boldsymbol{\Lambda}_i = \text{diag}[\sigma_{i,1}^2, \dots, \sigma_{i,d}^2]$ 分别表示第 i 个回归子中核函数的中心向量和尺度矩阵.

小波具有较好的时频局部化性能, 在时频分析、模式识别中也有广泛的应用, 采用如下形式的 d 维小波函数:

$$\phi_i^{(W)}(\mathbf{x}(t); \boldsymbol{\sigma}_i, \mathbf{c}_i) = \exp\left\{-\frac{1}{2}[\mathbf{x}(t) - \mathbf{c}_i]^T \boldsymbol{\Lambda}_i^{-1}[\mathbf{x}(t) - \mathbf{c}_i]\right\} \cdot \prod_{j=1}^d \cos\left(1.75 \cdot \frac{x_j(t) - c_{ij}}{\sigma_{ij}}\right)$$

常用的核函数还有多项式核、Sigmoid 核、B-样条核. 本文将高斯核和小波核函数结合起来, 构造如下形式的混合核函数:

$$\phi = a\phi^{(W)} + (1-a)\phi^{(G)}, \quad 0 \leq a \leq 1$$

a 是调节小波核函数与高斯核函数的参数. 显然, 当 $a = 0$ 时, 混合核函数即为高斯核函数; 当 $a = 1$ 时, 混合核函数退化为小波核函数. 通过调节参数 a , 使混合核函数适应不同的数据类型, 等效于在核函数的选择中, 融入对具体问题的先验知识. 图 2 为一维混合核函数当 a 分别取 0.2, 0.4, 0.7, 0, 1 时的曲线图, 其中核中心 $c = 1$, 尺度 $\sigma = 1$. 从图中可以看出混合核函数兼具高斯核函数和小波核函数的特性.

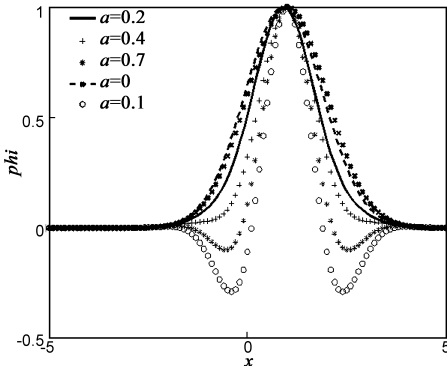


图2 混合核函数曲线

3.2 核函数中心

当样本长度不是很长时, 可以考虑用样本数据作为核函数中心的备选项. 当样本长度较大时, 可以用 k 均值聚类的方法得到聚类中心, 作为核函数中心备选项.

k 均值聚类要预先知道聚类个数, 本文利用 Krzanowski 和 Lai 提出的 sum-of-squares 准则得到最佳的个数 N_c ^[10]. 则混合核 RBF 神经网络的核函数中心备选项有 N_c 个.

由于传统的 k 均值聚类算法对初始值的选取非常敏感, 导致聚类的结果往往只是局部最优值. 为了克服这个问题, 本文将采用改进的全局 k 均值算法得到核函数中心的备选项^[11], 具体算法如下:

矩阵 \mathbf{X} 由 N 个待聚类的样本组成, $\mathbf{X} = [\mathbf{z}_1, \dots, \mathbf{z}_N]^T$, 其中, 每个样本均为 p 维, $\mathbf{z}_i = [z_{i1}, \dots, z_{ip}]$.

Step 1 进行初始化, 计算 \mathbf{X} 的质心 \mathbf{x}^1 :

$$\mathbf{x}^1 = \frac{1}{N} \sum_{i=1}^N \mathbf{z}_i, \quad \mathbf{z}_i \in \mathbf{X}, \quad i = 1, \dots, N$$

并且置 $q = 1$;

Step 2 令 $q = q + 1$, 如果 $q > N_c$, 即终止;

Step 3 考虑第 q 个聚类中心时, 将之前得到的聚类中心 $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^{q-1}$, 连同 \mathbf{X} 中的每一个点 \mathbf{z} , 共 N 个初始向量集合 $W_i = \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^{q-1}, \mathbf{z}_i\}$, $i = 1, \dots, N$ 分别进行 k 均值聚类, 最佳的聚类结果保留下来, 记为 $\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^q$;

Step 4 置 $\mathbf{x}^i = \mathbf{y}^i$, $i = 1, \dots, q$, 返回 Step 2.

3.3 核函数尺度

采用可调的尺度构造 RBF 网络. 在每一个寻求神经元的步骤中, 最小化均方误差(5). 具体一点, 对每一个聚类中心, 运用搜索算法找到相应的“最优”尺度, 使均方误差最小的(中心, 尺度, 混合核调节参数)组合即为第一个回归子核函数参数, 如此循环, 直至均方误差达到阈值或到达最大迭代次数. 许多随机搜索算法都可以用来决定第 L 步回归子的参数, 例如遗传算法, 模拟退火算法等.

首先, 介绍 GSO 算法的搜索角度的表示方法. 在一个 n 维搜索空间中, 第 k 代中的第 i 个元素具有当前位置值(鱼所在的位置) $\boldsymbol{\sigma}_i^k \in R^n$, 和一个朝向角度(鱼游的方向) $\boldsymbol{\varphi}_i^k = (\varphi_{i1}^k, \dots, \varphi_{i(n-1)}^k) \in R^{n-1}$, 搜索的方向向量 $D(\boldsymbol{\varphi}_i^k) = (d_{i1}^k, \dots, d_{in}^k) \in R^n$ 可由朝向角度计算出来:

$$d_{i1}^k = \prod_{q=1}^{n-1} \cos(\boldsymbol{\varphi}_i^k)$$

$$d_{ij}^k = \sin(\boldsymbol{\varphi}_{i(j-1)}^k) \prod_{q=j}^{n-1} \cos(\boldsymbol{\varphi}_i^k), \quad j = 2, \dots, n-2$$

$$d_{in}^k = \sin(\boldsymbol{\varphi}_{i(n-1)}^k)$$

在优化 RBF 网络的第 L 项时, 利用如下的 GSO 算法:

置 $k = 0$, 随机初始化尺度和混合核调节参数的位置值 $\boldsymbol{\sigma}_i$ 和朝向角度 $\boldsymbol{\varphi}_i$, $i = 1, \dots, P_s$, 组成种群, P_s 指的是种群的规模. 利用 3.2 节得到的聚类中心集合 $\mathbf{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_{N_c}\}$, 计算 \mathbf{c}_{k+1} 和 $\boldsymbol{\sigma}_i$. 根据式(5)计算均方误差 $f_{c_{k+1}}(\boldsymbol{\sigma}_i)$. 固定 \mathbf{c}_{k+1} , 做以下循环:

While 停止条件不满足

For 种群中的每个元素

选择复制器: 选择均方误差对应的尺度和混合核调节参数作为复制器, 记为 $\boldsymbol{\sigma}_p^k$.

开始复制:

(1) 零角度搜索: $\boldsymbol{\sigma}_z = \boldsymbol{\sigma}_p^k + r_1 l_{\max} D_p^k(\boldsymbol{\varphi}^k)$;

(2) 右搜索:

$\boldsymbol{\sigma}_r = \boldsymbol{\sigma}_p^k + r_1 l_{\max} D_p^k(\boldsymbol{\varphi}^k + r_2 \theta_{\max}/2)$;

(3) 左搜索:

$\boldsymbol{\sigma}_l = \boldsymbol{\sigma}_p^k + r_1 l_{\max} D_p^k(\boldsymbol{\varphi}^k - r_2 \theta_{\max}/2)$;

其中 $r_1 \in R^1$ 是一个均值为 0, 方差为 1 的正态分布随机变量, $r_2 \in R^{n-1}$ 是正方体 $[0, 1]^{n-1}$ 中均匀分布的随机向量. $l_{\max} = \sqrt{n}$ 代表事先

选定的搜索范围. 如果 $\sigma_x, \sigma_y, \sigma_z$ 的适应值小于 σ_p , 则将复制器的位置值更换为最小的. 否则位置值不变, 更换复制器的朝向角度 ϕ^k . $\phi^{k+1} = \phi^k + r_2 \alpha_{\max}$, 其中 α_{\max} 是最大旋转角度. 如果 a 次循环后, 复制器仍然没有找到更好的角度, 则 $\phi^{k+a} = \phi^k$;

开始搜寻: 在剩余种群中随机寻找 80% 元素进行搜寻, 搜寻元素将向目前最优元素搜寻, 即: $\sigma_i^{k+1} = \sigma_i^k + r_3 \cdot (\sigma_p^k - \sigma_i^k)$. 这里 \cdot 指的是向量之间的 Hadamard 积; $r_3 \in R^n$ 是产生于 $(0, 1)$ 的均匀分布随机序列.

开始分散: 对剩余种群进行分散处理, 分散方式如下:

(1) 角度分散:

$$\phi^{k+1} = \phi^k + r_2 \alpha_{\max};$$

(2) 位置值分散: 产生随机距离 $l_i = a \cdot r_1 l_{\max}$, 然后产生随机位置值:

$$\sigma_i^{k+1} = \sigma_i^k + l_i \cdot D_i^k(\phi^{k+1});$$

计算新种群的适应度值

End for

$k = k + 1;$

End while

将最小化均方误差的中心、尺度和混合核调节参数的组合作为 RBF 网络的第 L 项. 当误差或项数到达阈值时停止搜索新的神经元.

4 算法

混合核 RBF 网络的构造算法如下:

Step1 计算出最佳聚类个数 N_c ; 利用全局 k 均值聚类算法计算出中心 c_1, \dots, c_{N_c} , 作为核函数中心候选项;

Step2 将尺度 σ_i 和混合核调节参数 a 组成一个向量 u , 对每一个聚类中心 c_1, \dots, c_{N_c} , 用 GSO 算法搜索使均方误差最小的 u , N_c 个均方误差中最小的对应中心和向量 u 即为第一个回归子参数. 当用 GSO 搜索第 L 个回归子时, 可由 3.3 中方法进行.

Step3 当逐项回归的误差到达噪声的方差(可预先估计)时停止回归, 或者也可以利用某些准则(例如 generalized cross-validation 准则^[3])来确定误差的阈值.

5 实验

RBF 网络模型的稀疏性, 通常可以用如下的 $l_p (p \leq 1)$ 模来作为度量:

$$\|X\|_p = \left(\sum_{j=1}^n |x_j|^p \right)^{\frac{1}{p}}$$

如果一个信号有稀疏表达, 则 $l_p (p \leq 1)$ 达到一个比较小的值. 同时稀疏性可以用一种直观的方式来衡量: 设 X 和 Y 两个序列, 将它们各自的分量按绝对值进行从大到小的顺序排列后, 可以认为向零值衰减快的序列比另一个序列具有更稀疏的表达形式.

实验均在 Matlab 上进行. GSO 算法的种群规模为 48, 方向角初值为 $\phi^0 = (\pi/4, \dots, \pi/4) \in R^{n-1}$, 复制操作

循环次数 $a = \text{round}(\sqrt{n+1})$, round 指的是取最近整数函数, n 为搜索空间的维数. 最大旋转角度 $\theta_{\max} = \alpha_{\max} = \pi/a^2$.

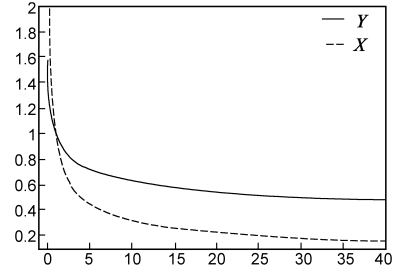


图3 X序列比Y序列表示更稀疏

实验 1 一维数据仿真

考虑如下的 sinc 函数:

$$f(x) = \frac{\sin(x)}{x}, \quad -10 \leq x \leq 10$$

产生 400 个等间隔的带噪声数据 $y = f(x) + \epsilon$, 其中, ϵ 是零均值、标准差为 0.2 的高斯白噪声. 200 个偶数点用于训练, 剩下的 200 个奇数点用于测试. 采用如下的混合核核函数:

$$\phi = a\phi^{(W)} + (1-a)\phi^{(G)}, \quad 0 \leq a \leq 1$$

$$\phi_m^{(G)}(x) = \exp\left(-\frac{(x-c_m)^2}{2\sigma_m^2}\right)$$

$$\phi_m^{(W)}(x) = \exp\left(-\frac{(x-c_m)^2}{2\sigma_m^2}\right) \cdot \cos\left(1.75 \cdot \frac{(x-c_m)}{\sigma_m}\right)$$

G, W 分别表示高斯核和小波核, c_m, σ_m 分别是核函数中心和尺度.

运用 sum-of-squares 准则, 得出最佳的聚类个数 $N_c = 5$. 利用全局 k 均值聚类得到中心. 对每一个中心, 运用 GSO 搜索算法搜索最佳的尺度 σ 和混合核调节参数 a , 均方误差最小的组合即为第一个回归子参数, 接着, 利用 OLS 对误差进行逐项回归分析, 训练 RBF 网络.

图 4 为 Billings 等人提出的 MSRBF 对该数据训练并拟和的结果. 训练的 RBF 网络由 8 项高斯核函数的线性加权得到. 训练误差为 0.0384, 测试误差为 0.0396^[7]. 图 5 是本文提出的混合核 RBF 网络拟和的结果. 训练的 RBF 网络由 4 项混合核函数的线性加权得到, 训练

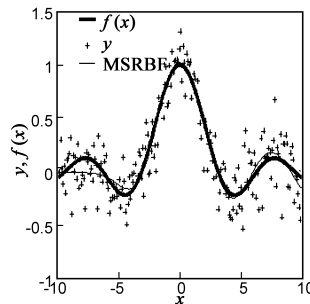


图4 MSRBF拟和结果

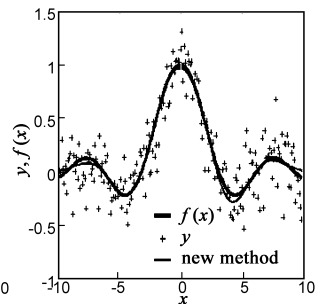


图5 混合核RBF拟和结果

误差为 0.0376,测试误差为 0.0381.每个回归子的尺度、中心、混合核调节参数和权值具体结果见表 1.同样的实验数据运用在 2005 年 Chen. S 提出的可调核的正交最小二乘回归方法(OLSTK)中,网络由 7 项高斯核叠加得到.训练误差是 0.0394,测试误差是 0.0390^[3].

表 1 混合 RBF 网络模型

Centre	Scale	Weight	a
-0.5578	3.0531	2.3067	0.7836
-3.5847	3.4107	0.1359	1.4603
1.9424	2.3990	0.5763	1.1057
-0.5578	4.1330	-1.2138	0.8642

图 6 为 MSRBF、新的混合核 RBF 以及 OLSTK 所得网络的稀疏度量图.从图 6 可以看出,相比与其他两种方法,新方法的稀疏有较明显的下降,这说明了新的 RBF 网络更加稀疏.

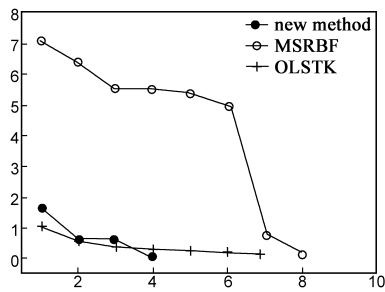


图 6 MSRBF、OLSTK和混合RBF网络的稀疏度量图

实验 2 三维实际数据处理

本实验中将对 Leyland TL11 中燃油消耗^[12](输入 u_k ,图 7 所示)和发动机转速(输出 y_k ,图 8 所示)的关系进行非线性系统建模,构造 RBF 神经网络.410 个数据点中,前 210 个用于训练,后 200 个用于测试.已有的研究表明^[12],这个数据满足模型: $y_k = f(x_k) + e_k$,其中, $x_k = [y_{k-1}, u_{k-1}, u_{k-2}]^T$.

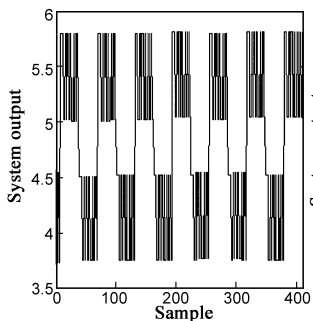


图 7 系统输入 u_k

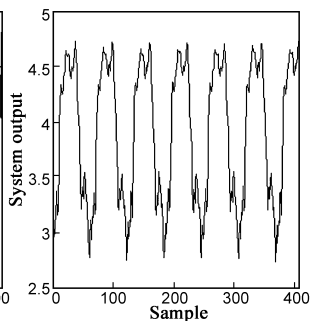


图 8 系统输出 y_k

图 9 为可调尺度的混合核 RBF 网络预测的结果. RBF 网络由 9 个混合核的线性组合得到,每项的尺度、中心、混合核调节参数和权值见表 2,训练误差为 0.0005,测试误差为 0.0005.图 10 为模型真实输出与预测输出的误差图,可以看出,混合核 RBF 网络预测误差

在 ± 0.1 左右.文献[12]对同样的数据得到了 11 项高斯核的 RBF 网络,训练误差为 0.000496,测试误差为 0.000503.

表 2 可调尺度的混合核 RBF 网络模型

Centre		Scale		a	Weight		
3.8977	4.7992	4.7836	4.4379	3.8308	4.0018	0.1107	9.3863
1.7288	1.8720	2.2600	2.0253	3.7455	3.7594	0.0737	-20.5060
1.4309	1.8817	1.8817	1.8370	1.5543	1.5711	0.4535	-8.3136
1.4424	1.8720	1.8720	1.4329	1.0820	0.7458	0.3215	-56.6647
3.8977	4.7992	4.7836	2.8038	2.7225	2.8150	0.4435	-2.0979
1.4424	1.8720	1.8720	2.3161	2.9889	3.4351	0.0739	28.7580
1.4309	1.8817	1.8817	4.9333	3.3820	4.1037	0.0780	-7.8459
3.8977	4.7992	4.7836	0.6784	4.1176	2.4196	0.0619	-0.2339
1.6982	1.8720	1.8720	3.8080	0.6649	0.6658	0.0005	646.0583
Training MSE				0.0005			
Testing MSE				0.0005			

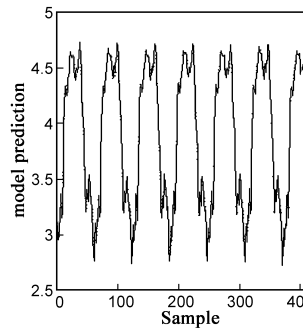


图 9 模型预测图

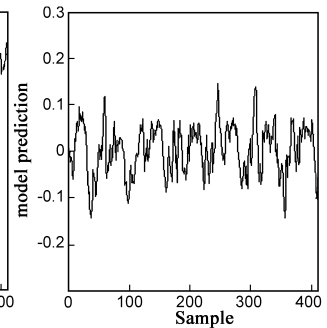


图 10 误差图

6 结论

为了得到更加稀疏的 RBF 网络,本文提出了一种尺度可调的混合核 RBF 网络构造方法.利用高斯核函数和小波核函数的特点,将两者的线性组合构成混合核,通过混合核调节参数得到不同性质的核函数.通过群搜索优化器搜索算法对尺度以及混合核调节参数进行最优搜索,并利用 OLS 逐项回归分析.实验结果说明,新的混合核 RBF 网络比传统单核网络更稀疏.当然,也可以考虑用别的核函数的线性组合.如何更好的根据数据的分布特点选择适合的混合核函数模型,这将是下一步要研究的内容.

参考文献:

[1] 臧小刚,宫新保,等.一种基于免疫系统的 RBF 网络在线训练方法[J].电子学报,2008,36(7):1396-1400.
Zang xiao-gang, Gong xin-bao, et al. An online RBF network based on immune system[J]. Acta Electronica Sinica, 2008, 36(7):1396-1400. (in Chinese)

[2] Chen S, Donoho D L, Saunders M. Atomic decomposition by basis pursuit[J]. SIAM J. Sci Comp, 1999, 20(1):33-61.

[3] Chen S, Wang X X, Brown D J. Orthogonal least squares re-

- gression with tunable kernels[J]. Electron Lett 2005, 41(8): 484 - 486.
- [4] Karayiannis N B. Reformulated radial basis neural networks trained by gradient descent[J]. IEEE Trans on Neural Networks, 1999, 10(3): 657 - 671.
- [5] Lazaro M, Santamaria I, Pataleon C. A new EM-based training algorithm for RBF networks[J]. Neural Networks, 2003, 16(1): 69 - 77.
- [6] Gonzalez J, Rojas I, Ortega J, et al. Multi-objective evolutionary optimization of the size, shape and position parameters of radial basis function networks for function approximation[J]. IEEE Trans on Neural Networks, 2003, 14(6): 1478 - 1495.
- [7] Billings S A, Wei H, Balikhin M. Generalized multi-scale radial basis function networks[J]. Neural Networks, 2007, 20(10): 1081 - 1094.
- [8] Tan Y, Wang J. A support vector machine with a hybrid kernel and minimal Vapnik Chervonenkis Dimension[J]. IEEE Trans on Knowledge and Data Engineering, 2004, 16(4): 385 - 395.
- [9] He S, Wu Q H, Saunders J R. Group search optimizer: an optimization algorithm inspired by animal searching behavior[J]. IEEE Trans on Evolutionary Computation, 2009, 13(5): 973 - 990.
- [10] Krzanowski W J, Lai Y T. A criterion for determining the number of groups in a data set using sum of-squares clustering [J]. Biometrics, 1988, 44(1): 23 - 34.
- [11] Likas A, Vlassis M, et al. The global k-means clustering algorithm[J]. Pattern Recognition, 2003, 36(2): 451 - 461.
- [12] Chen S, Wang X X, Harris C J. NARX-based nonlinear system identification using orthogonal least squares basis hunting

[J]. IEEE Trans on Control Systems Technology, 2008, 16(1): 78 - 84.

作者简介:



付丽华 女, 1979 年生于湖北枝江, 博士, 中国地质大学数理学院(武汉)副教授. 研究方向为模式识别、信号处理.

E-mail: fulihua9270@yahoo.com.cn



李宏伟 男, 1965 年生于湖南汨罗, 博士, 中国地质大学数理学院(武汉)教授, 博士生导师. 研究方向为信号处理、机器学习等.

E-mail: hwli@cug.edu.cn



张猛 男, 1977 年出生于湖北武汉, 博士, 华中师范大学副教授, 硕士生导师. 研究方向为数据挖掘、模式识别等.

E-mail: morosezhang@yahoo.com.cn