

径向基函数网络结构确定的新算法^①

李波^②

(天津大学管理学院)

王秀峰

(南开大学计算机与系统科学系)

【摘要】讨论了对于径向基函数网络的一种结构确定双向回归新算法. 在这种算法中, 采用正交化手段和“新息—贡献”准则, 可同时确定径向基函数网络的结构(即中心点)和参数估计, 不仅克服了传统算法的从一大组输入数据中任意选择中心点常出现的数值病态和选择结果依赖于数据排列顺序等问题, 而且大大减少了计算量且保证能得到唯一的最优网络描述. 最后用仿真例子说明了这种新算法的正确性和有效性.

关键词: 径向基函数, 神经网络, 建模

分类号: O224, F832.5

TP18

0 引言

目前, 径向基函数网络(the radial basis function, RBF) 在信号处理等许多领域中得到越来越广泛的应用^[1~4], 这主要是因为 RBF 网络和神经网络之间具有密切关系. 把 RBF 网络看作是一个固定其所有中心点和隐含层的非线性特性的、参数线性的特殊的两层神经网络, 其中它的隐含层起到了固定非线性转换作用, 它把输入空间映射到一个新空间上去, 而输出层则实现这个新空间上的线性组合作用, 且需要调节的参数只是这个线性组合器的权数. 理论证明只要 RBF 网络的隐含层适当固定, RBF 网络具有和神经网络类似的近似特性^[5,6].

理论研究和实验表明在 RBF 网络中, 非线性函数的选择对于网络性能的影响并不非常重要, 重要的是 RBF 网络中所选的中心点, 过去许多的研究常简单地假设中心点从一大组输入数据点中任意选取, 这样构造的 RBF 网络要么性能不好, 要么有较大的维数. 后来较多地采用聚类的方法实现中心点的选择, 但此法对数据比较敏

感^[7]. 文[8]提出了一种正交最小二乘选择算法, 该方法可同时确定中心点和输出数值, 但存在一个很大缺点是中心点的选择强烈地依赖数据排列顺序, 因此造成对同一组数据所选出的中心点数目不唯一, 且所选择的中心点数很可能不是最少的.

本文借助于在文献[9]提出的新息—贡献准则对中心点的选择, 提出正交优选的双向回归算法, 不仅避免了随机选择中心点时常发生的数值病态问题, 而且所选中心点不依赖于数据排列的顺序, 可唯一得到一最简的 RBF 网络. 整个选择过程递推进行, 可同时确定中心点(即网络结构), 大大减少了计算量.

1 径向基函数网络

考虑具有 n 个输入和一个输出的 RBF 网络, 如图 1 所示, 网络实现的映射关系为 $f_r: R^n \rightarrow R$.

$$f_r(X) = \lambda_0 + \sum_{i=1}^n \lambda_i \varphi(\|X - C_i\|) \quad (1)$$

① 国家自然科学基金重大项目(79790130)和天津自然科学基金(993600811)资助项目.

② 李波, 硕士, 讲师. 通讯地址: 天津大学管理学院, 邮编: 300072.
本文 1998 年 7 月 20 日收到.

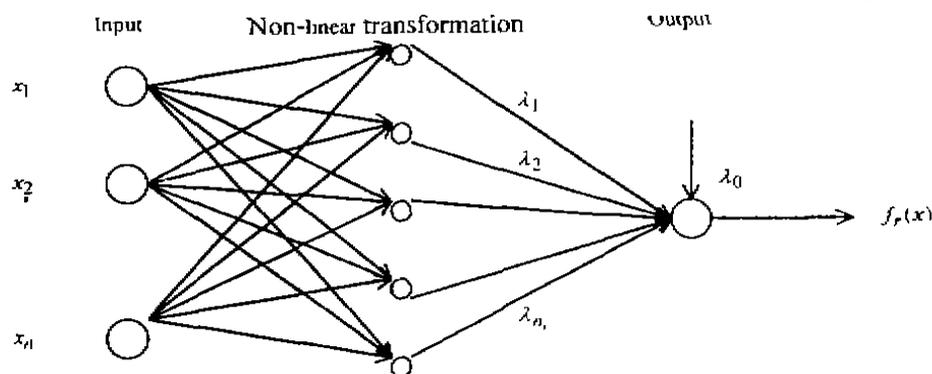


图 1 RBF 网络的结构

其中 $X \in R^n$ 是输入向量, $\varphi(\cdot): R^n \rightarrow R$ 是给定非线性函数, $\|\cdot\|$ 表示欧几里得 (Euclidean) 范数, $\lambda_i, 0 \leq i \leq n$ 是加权, $C_i \in R^n, 1 \leq i \leq n$ 是 RBF 中心点, n 是中心点的个数.

在 RBF 网络中, 非线性函数 $\varphi(\cdot)$ 和中心点 C_i 一般假设是固定的, 取得 N 组输入 $X(t)$ 和相应的期望输出 $y(t), t = 1, \dots, N$, 加权 λ_i 由线性最小二乘法可进行估计, 但是 $\varphi(\cdot)$ 和 C_i 的选择须慎重, 否则 RBF 网络不能拟合两层神经网络的性能. 理论研究和实际结果表明^[9,11] 非线性函数 $\varphi(\cdot)$ 的选择对 RBF 网络的性能不很重要, 典型的非线性函数 $\varphi(\cdot)$ 常取作

薄板样条函数 $\varphi(v) = v^2 \log(v)$ (2)

高斯函数 $\varphi(v) = \exp(-v^2/\beta^2)$ (3)

其中 β 是一实常数

常用的还有其它两种函数

$\varphi(v) = (v^2 + \beta^2)^{1/2}$ (4)

$\varphi(v) = (v^2 + \beta^2)^{-1/2}$ (5)

对固定的中心点 C_i 和给定的非线性函数 $\varphi(\cdot)$, 则 $\varphi(\|X - C_i\|)$ 可用符号 P_i 表示, 显然 $P_i(t) = P_i(X(t))$ 是 $X(t)$ 的函数, 这时 RBF 网络式(1) 看作为线性回归模型的特殊情形.

$$y(t) = \sum_{i=1}^n \theta_i P_i(t) + \epsilon(t) \quad (6)$$

其中 $y(t)$ 是期望输出, θ_i 是参数, $P_i(t)$ 是回归项, $\epsilon(t)$ 是误差, 与回归因子 $P_i(t)$ 不相关. 这样就选择 RBF 网络的适当中心点问题转化为从给定的一大组回归量中选择有意义的子集问题.

2 正交最小二乘选择方法^[6]

假设中心点的数目为输入的样本数, 隐含单元函数的中心值也取输入样本的数值. 取得 N 组数据 $\{Y(t), P(t), t = 1, \dots, N\}$, 则式(6) 的矩阵形式为

$$Y = P \cdot \theta + \epsilon \quad (7)$$

其中 $Y = [y(1), \dots, y(N)]^T, \theta = [\theta_1, \dots, \theta_n]^T$

$$\epsilon = [\epsilon(1), \dots, \epsilon(N)]^T \quad P = [P_1, \dots, P_n]$$

$$P_i = [P_i(1), \dots, P_i(N)]^T \quad 1 \leq i \leq n$$

把矩阵 P 进行正交化分解

$$P = V \cdot A \quad (8)$$

$$\text{其中 } A = \begin{bmatrix} 1 & a_{12} & \dots & a_{1n} \\ 0 & 1 & a_{23} & \dots & a_{2n} \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & & & 1 \end{bmatrix}_{n \times n}$$

V 是一个 $N \cdot n$ 具有列 V_i 正交的矩阵.

由此得到式(7) 的辅助模型 $Y = Vg + \epsilon$ (9)

参数估计为 $g = (V^T V)^{-1} V^T Y$ (10)

$$\text{即 } g_i = \frac{\sum_{t=1}^N v_i(t) y(t)}{\sum_{t=1}^N v_i^2(t)} \quad i = 1, \dots, n$$

具体地由古典 Gram-schmidt 方法,

$$\begin{cases} v_1 = P_1 \\ \alpha_i = v_i^T P_k / (v_i^T v_i), 1 \leq i \leq j \\ v_j = P_j - \sum_{i=1}^{j-1} \alpha_i v_i, j = 2, \dots, n \end{cases} \quad (11)$$

而原模型式(7) 中的参数 θ 满足

$$A\theta = g \quad (12)$$

这样通过正交化过程同时确定了 RBF 网络的中心点和输出数值.

3 正交优选双向回归的递推算法

从文献[8]所给算法不难看出,所选中心点数目及中心点的值强烈地依赖于数据排列顺序,如果换一种排列顺序,则很可能产生完全不同的结果,即中心点的数目和数值都将会不同.为此提出基于“新息—贡献”准则的双向回归递推算法,不仅具有正交算法可克服病态的优点,而且可以克服不唯一性这一重大缺点,可得到唯一最佳网络结构.

3.1 “新息—贡献”准则^[9]

假设 N 维向量 P_1, \dots, P_k , 相应地正交化向量为 $\tilde{v}_1, \dots, \tilde{v}_k$, 记 $R_k = [\tilde{v}_1, \dots, \tilde{v}_k]$, 对 $\forall P_i$, 则 P_i 可由 $\tilde{v}_1, \dots, \tilde{v}_k$ 线性表出的部分为

$$u = R_k(R_k^T R_k)^{-1} R_k^T P_i \quad (13)$$

P_i 所带来的“新息”为

$$E = P_i - u = [e(1), \dots, e(N)] \quad (14)$$

P_i 对于描述 Y 上的“贡献”为

$$\mathcal{E} = \frac{E^T Y}{E^T E} = \frac{\sum_{t=1}^N e(t)y(t)}{\sum_{t=1}^N e^2(t)} \quad (15)$$

“新息—贡献”准则是说衡量一个变量是否重要,不仅看它带来多少“新息”,而且看该“新息”对描述输出的重要程度.将式(13)、(14)展开,并同式(11)比较易见, E 恰好是 P_i 与 $\tilde{v}_1, \dots, \tilde{v}_k$ 的正交化向量,因此正交化和新息值的计算是一致的,可大大减少计算量.

3.2 双向回归递推算法

该算法仍基于模型(7),对(7)中每一回归项按“新息—贡献”准则逐个选取,并判断是否加入模型.基本思想是:首先从所有回归项中选择“贡献”最大的项加入回归模型,以后每一步在剩余的项中选出“贡献”最大的项,若在统计意义下,该项对输出精度提高的影响是显著的,则被选入模型.若由于新增加的项使得某个已选入的项变得不重要了,则将该项剔除,然后再从剩余项中选“贡献”最大的项,重复上述步骤,直到既无可选入项又无可剔除项为止.

具体计算过程中,采用了文献[12]中给出的运算技巧,实现过程递推进行,每一步计算量相当少.

例如 $R_1 = [P_1], \Phi_1 = [R_1^T R_1], \Phi_1^{-1} = [R_1^T R_1]^{-1} = 1/P_1^T P_1$.

在逐步选项过程中,如第 $K+1$ 步

$$R_{k+1} = [R_k, \tilde{v}_{k+1}]$$

$$\text{则 } \Phi_{k+1} = [R_{k+1}^T R_{k+1}]^{-1} = \begin{bmatrix} \Phi_k^{-1} & 0 \\ 0 & 1/\sum_{t=1}^n \tilde{v}_{k+1}^2(t) \end{bmatrix}$$

对剔除某项,可简单地实现,详见文[9]和[12].这样就实现了递推计算,正交化性质和递推计算使得整个选项过程简单且效率高.

平衡最终网络复杂度和精度的重要参数是进行模型检验.在许多信号处理应用中,常用的是误差减少率^[6]和 AIC 检验,因为误差减少率仅强调网络的性能,在算法中采用了折衷网络性能和参数个数(涉及网络复杂度)的 AIC 检验和 F 检验,如在选项的第 K 步:

$$\text{定义 } E_k = Y - \sum_{i=1}^k g_i \tilde{v}_i \quad (16)$$

$$S_k^2 = \frac{1}{N} E_k^T E_k \quad (17)$$

$$BIC_k = N \ln S_k^2 + k \ln N \quad (18)$$

$$F = (S_{k-1}^2 - S_k^2)/S_k^2 * (N - K) \quad (19)$$

每选入一向量,应同时检查 BIC 和 F 的值,以判断是否把该项选入模型,同时当每选入一项时,还要检验若从前所有选入项中删去某一项时对模型的影响,以避免某项是其它若干选入项的线性组合,这样不仅可以避免数值病态问题,而且可得到唯一最简单的 RBF 网络.

正交优选算法的选项过程总结如下:

(1) 从初始数据构造式(7)中的各向量 Y, P_i , 并且给出 F 检验的显著性指标 F_α .

(2) 第 1 步,计算每一 P_i ($i = 1, \dots, n$) 对 Y 的贡献,记最大贡献的相应项 P_{max} 为 \tilde{v}_1 ,基于式(10), (16), (17), (18), (19) 计算 g, E_1, S_1^2, BIC_1 和 F .

(3) 第 K 步 ($K \geq 2$) 基于“新息—贡献”的式(13), (14), (15) 计算其余各项对 Y 的贡献,取贡献最大的项记为 \tilde{v}_k ; 同样计算 g, E_k, S_k^2, BIC_k 和 F .

(4) 检验 BIC_k 和 F 准则值,确定是否把 \tilde{v}_k 选入模型或算法结束.

(5) 第 K 步 ($K \geq 2$), 对于(3), (4) 步选入模型中的 \tilde{v}_k , 检验其对前选入项 \tilde{v}_i ($i = 2, \dots, K$) 的

影响,去掉某一项 $\tilde{v}_i (i = 2, \dots, K)$, 分别重新计算 $E_{K(t)}$, $S_{K(t)}$, $BIC_{K(t)}$ 和 F , 把 $MINF$ 与 $BIC_{K(t)}$ 和 F 与 BIC_K 作比较, 确定是否应剔除某项 \tilde{v}_i 或算法结束.

4 仿真实例

用本文给出的算法进行股票交易市场的 RBF 网络建模研究, 并与文[8]中给出的算法进行了比较. 采集的数据取自 1995 年沪市青岛海尔一年中每天的平均价格(共 223 个数据), 令 $y(t)$ 是 t 时刻的时间序列值, 使用 RBF 网络进行建模即为 $\hat{y}(t) = f_c(x(t))$, 其中输入 $x(t) = [y(t-1), \dots, y(t-n_y)]^T$ 是序列的过去观察值.

由股票波动的规律, 取时滞 $n_y = 30$, 并且取 RBF 网络中的非线性变换 $\varphi(\cdot)$ 为高斯函数, 其中 $\beta^2 = 2$. 仿真过程中把数据集分成两部分, 前半年的 100 多数据点作为训练集进行建模, 后半年 100

多数据点作为测试集进行预测, 用本文给出的算法得到的最终 RBF 网络模型(只有 7 个隐含结点)为

$$\hat{y}(t) = \hat{\lambda}_0 + \sum_{i=1}^7 \hat{\lambda}_i \varphi(\|x(t) - c_i\|)$$

实测值 $y(t)$ 与拟合值 $\hat{y}(t)$ 的比较由图 2 所示, 误差在图 3 中表示; 用后半年的测试集数据进行测试得到的预测值和实测值的比较如图 4 所示. 由于股票数据的波动性及数据量不够大等原因, 影响了拟合的精度. 但如果采用文献[8]的方法对此实例同样进行建模研究, 辨识出的最终 RBF 网络(有 10 个中心点)模型为

$$y(t) = \lambda_0 + \sum_{i=1}^{10} \lambda_i \varphi(\|x(t) - c_i\|)$$

预测值 $\hat{y}(t)$ 和实测值 $y(t)$ 的比较由图 5 所示. 由此可以看出, 本文算法比文献[8]算法选出的中心点个数要少且有较高的精度, 可得到最简单的 RBF 网络.

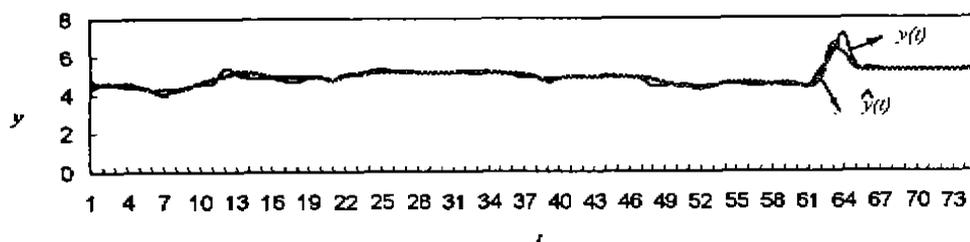


图 2 本文算法实测值与拟合值比较

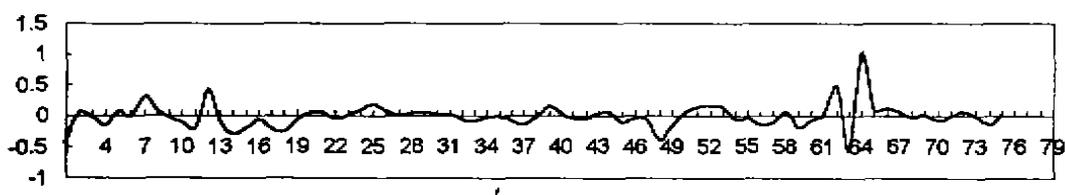


图 3 本文算法的误差序列

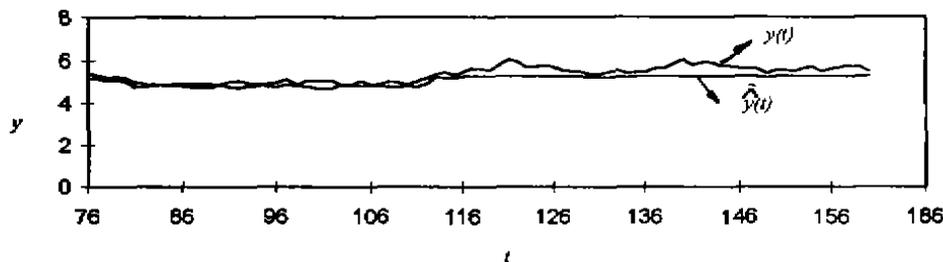


图 4 测试集数据的实测值与预测值比较

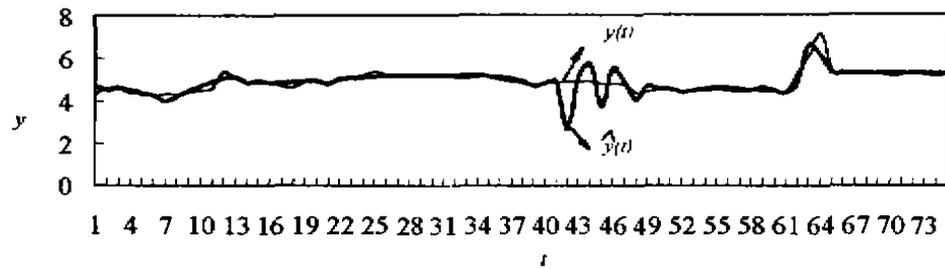


图5 文献[8]的实测值与拟合值比较

参 考 文 献

- 1 Chen S, Billings S A, Cowan C F N, et al. Nonlinear systems identification using radial basis functions. *Int. J. Syst. Sci.*, 1990;21(12):2513~2539
- 2 Chen S, Billings S A, Cowan C F N, et al. Practical identification of NARMAX models using radial basis function. *Int. J. Contr.*, 1990;52(6):1327~1350
- 3 Bruce A, Timothy D C. Cooperative-competitive genetic evolution of radial basis function centers and widths for time series prediction. *IEEE Tran. on Neural Networks*, 1996;7(4):869~880
- 4 Michael H, Stefan K. Multiplication-free radial basis function network. *IEEE Tran. on Neural Networks*, 1996;7(6):1461~1464
- 5 Broomhead D S, Lowe D. Multivariable functional interpolation and adaptive networks. *Complex System*, 1988, 2:321~355
- 6 Girosi. Some extensions of radial basis functions and their applications in artificial intelligence. *Computers Math. Applic.*, 1992;24(12):61~80
- 7 Kohonen T. An introduction to neural computing. *Neural Network*, 1988;1(1):3~16
- 8 Chen S, Cowan C F N, Grant P M. Orthogonal least squares learning algorithm for radial basis function network. *IEEE Transactions on Neural Networks*, 1991;2(2):302~309
- 9 Wang Xiufeng, Li Bo. Selforganizing methods modelling with "concentrated" variables. *System Analysis Modeling Simulation*, 1995;20:51~57
- 10 Powell M J D. Radial basis functions for multivariable interpolation: A review. *IMA Conf. Algorithms for the Approximation of Functions and Data*, New York; Oxford University Press, 1987. 143~167
- 11 Powell M J D. Radial basis function approximations to polynomials. *proc. 12th Biennial Numerical Analysis Conf. (Dundee)*, 1987. 223~241
- 12 王秀峰, 劳育红. 非线性动态系统模型结构确定和参数估计的新算法. *自动化学报*, 1992;18(4):385~392

A New Algorithm for The Structure Determination of Radial Basis Function Networks

Li Bo

School of Management, Tianjin University

Wang Xiufeng

Department of Computer & System Sciences, Nankai University

Abstract A stepwise regression new algorithm for the structure determination of radial basis function net-

works is proposed. In this algorithm, based on the orthogonalization method and the "Innovation-Contribution" criterion, the structure of radial basis function networks and parameter estimation are simultaneously determined. It not only avoids ill-conditioned problems occurred in many common learning algorithm which is often based on first choosing randomly some data points as radial basis function centers and then using singular value decomposition to solve for the weights of the network, but also can reduce considerably the amount of computations and we can obtain the optimal radial basis function networks. Finally, the identification of a simulated system show the correctness and effectiveness of this new algorithm.

Keywords: the radial basis function, neural networks, modeling

~~~~~  
(上接第 13 页)

- 2 陈国权,王 斌,陈玉祥.面向可持续发展的企业经营环境分析与经营管理框架,北京:清华大学学报,1998;4(13):58~65
- 3 Pollution Prevention Pays. 3M, July 1998

## Sustainable Development and the Future of Enterprise

*Chen Yuxiang*

The Administrative Center for China's Agenda 21

*Chen Guoquan*

School of Economics & Management, Tsinghua University

**Abstract** Based on the analysis of the changing business environment, it is pointed out in the paper that sustainable development has been becoming the basic element for improving the competitiveness of enterprise. The learning model for incorporating sustainable development into corporate management is put forward. From the perspectives of strategy formulation, technology selection, and production and marketing management, the basic way for the integrating philosophy of sustainable development with corporate management has been formulated.

**Keywords:** sustainable development, corporate management, competitiveness