

基于自适应投影学习算法的径向基函数网络设计及应用

张茁生, 刘贵忠, 刘 峰

(西安交通大学电子与信息工程学院信息与通信工程系, 西安 710049)

摘 要: 基于自适应投影原理, 本文提出了构造和训练径向基函数网络的自适应投影学习算法, 该算法能迭代地确定径向基函数的个数、中心的位置以及网络的权系数, 从而提供了生成径向基函数网络的一个简单和有效的方法. 通过应用于信道均衡器和时间序列预报信号处理领域中的两个例子, 进一步证明了该算法是行之有效的.

关键词: 径向基函数网络; 自适应投影算法; 结构设计; 信道均衡器; 时间序列预报

中图分类号: TP183 **文献标识码:** A **文章编号:** 0372-2112 (2000) 09-0120-03

Radial Basis Function Networks Based on Adaptive Projective Learning Algorithm and Its Applications

ZHANG Zhuosheng, LIU Guizhong, LIU Feng

(Department of Information & Communication Engineering, Xidian University, Xi'an 710049, China)

Abstract: In this paper, an adaptive learning procedure for constructing and training radial basis function networks is proposed, based on the adaptive projective algorithm. The procedure chooses the centers of radial basis function and the weights of networks one by one until an adequate network has been constructed, providing a simple and efficient means for growing radial basis function networks. This is illustrated by two examples on communication channel equalizer and the prediction of a time series.

Key words: radial basis function networks; adaptive projective algorithm; network structure; channel equalization; time series prediction

1 引言

神经网络作为一种新思想、新方法已广泛应用到信号处理、非线性系统辨识和机械人控制等领域中. 作为一种前馈型神经网络, 径向基函数(简记为RBF)网络目前越来越多地受到人们关注, 它含有单个隐层, 隐层结点的激活函数是径向基函数, 网络的输出是隐层结点输出的线性组合, 与多层前向神经网络相比, RBF网络以体积小、收敛快而见长, 并且其单隐层的结构十分有利于网络的分析 and 设计^[1]. 现已证明: 当隐层结点数充分多时, RBF网络可以任意精度地逼近 $L^2(R^n)$ 中的函数^[2]. RBF网络的性能依赖于径向基函数的数目及其中心的选择, 径向基函数的形状以及确定网络权系数所使用的方法. 然而, 理论研究和实际结果均表明: RBF网络的性能主要依赖于径向基函数的数目及其中心位置, 而径向基函数本身的选取不是至关重要的, 并且选取相同的结点宽度, 就可以任意地逼近 $L^2(R^n)$ 中的函数^[3]. 现存RBF网络的学习策略主要集中在三个方面, 即在径向基函数数目预先确定的条件下, RBF的中心从训练集中随机的选取或由训练集以某种无监督方式或有监督方式确定. 这就产生了象多层前向神经网络一样如何确定隐层结点数问题, 它常常导致网络具有较差的性能和较大的不必要的体积, 并且增加网络隐层结点数并不一

定能较好地改善网络的性能, 从而造成网络资源的浪费, 而RBF函数中心的选择显然也是不能令人满意的^[1,4].

为了克服上述缺点, 人们做了一些尝试, 主要是通过构造或删除的方法确定径向基函数的个数及其中心的位置, 或是开始使用一个结点, 然后根据需要增加新的结点或是开始使用很多结点, 然后当可能时合并它们^[4,5].

本文提出了一个新的构造和训练RBF网络的学习算法, 它建立在自适应投影原理的基础上, 充分地利用训练数据, 迭代地确定径向基函数的个数、中心的位置以及网络的权系数.

2 径向基函数网络

RBF网络是一个单隐层前馈型神经网络, 其结构如图1所示. 这个网络的输出为

$$y = X_0 + \sum_{i=1}^r X_i \phi(x - c_i) \quad (1)$$

其中, $x = (x_1, x_2, \dots, x_n)^T$ 是网络的输入向量; $\phi(\cdot)$ 是 R^n 中的欧几里得范数; $\phi(\cdot): R^+ \rightarrow R$ 径向基函数; X_i 是网络的权系数; c_i 是径向基函数的中心; r 是径向基函数的数目.

为了方便起见, 假定上述网络的输出只有一个. 由于多输出径向基函数网络可以分解成一组单输出的RBF网络, 结论可直接延拓到多输出情形. 为了构造一个RBF网络, 通常分

为两步: 一是要确定径向基函数的个数及其中心的位置; 二是确定网络的权系数. 前者是构造 RBF 网络的主要任务, 而后者是极小化一个二次函数, 可以较容易地求出其最小值点. 作为轴对称函数径向基函数 $\phi(\cdot)$ 是预先选定的, 通常选择高斯函数

$$G(x) = \exp(-x^2/R^2) \quad (2)$$

其中 R 确定了径向基函数的宽度.

3 R^N 中自适应投影算法及收敛性

基于 Gabor 信号集信号分解的自适应投影算法是 Mallat 在文[6]中创立的. 它是按最大投影匹配的准则寻找一组/最佳基, 以此达到信号的自适应分解的目的.

任取 $A = (a_1, a_2, \dots, a_N)^T, B = (b_1, b_2, \dots, b_N)^T \in R^N$, 则 R^N 中的内积和范数分别为

$$\langle A, B \rangle = \sum_{i=1}^N a_i b_i, \quad \|A\| = \sqrt{\langle A, A \rangle},$$

又设 $\{A^{(i)}\}_{i=1}^M \subset R^N, \langle A^{(i)}, A^{(i)} \rangle = 1$, 令

$$V = \text{span}\{A^{(i)} \mid 1 \leq i \leq N\}$$

则任取 $Y \in R^N$, 寻找 Y 在 V 上投影的自适应投影算法描述如下:

取 $A^{(i_0)} \in \{A^{(i)}\}_{i=1}^M$, 使

$$|\langle Y, A^{(i_0)} \rangle| = \sup_{1 \leq i \leq M} |\langle Y, A^{(i)} \rangle|$$

则 Y 可分解为 $Y = \langle Y, A^{(i_0)} \rangle A^{(i_0)} + R_Y^1$

易知, 残量 R_Y^1 与 $A^{(i_0)}$ 是正交的, 于是

$$\|Y\|^2 = \|\langle Y, A^{(i_0)} \rangle A^{(i_0)}\|^2 + \|R_Y^1\|^2$$

对残量 R_Y^1 进行同样分解, 如此反复, 第 n 次残量 R_Y^n 可分解为

$$R_Y^n = \langle R_Y^n, A^{(i_n)} \rangle A^{(i_n)} + R_Y^{n+1} \quad (3)$$

其中 $|\langle R_Y^n, A^{(i_n)} \rangle| = \sup_{1 \leq i \leq M} |\langle R_Y^n, A^{(i)} \rangle|$ (4)

于是, 在 p 次迭代后

$$Y = \sum_{k=0}^{p-1} \langle R_Y^k, A^{(i_k)} \rangle A^{(i_k)} + R_Y^p \quad (5)$$

$$\|Y\|^2 = \sum_{k=0}^{p-1} |\langle R_Y^k, A^{(i_k)} \rangle|^2 + \|R_Y^p\|^2 \quad (6)$$

其中 $R_Y^0 = Y$.

若记 P_V 为 R^N 到 V 的正交投影算子, Mallat 在文[6]中已证明, 当 $p \rightarrow \infty$ 时, R_Y^p 以指数速率收敛到 $Y - P_V Y$, 并且

$$P_V Y = \sum_{k=0}^{\infty} \langle R_Y^k, A^{(i_k)} \rangle A^{(i_k)} \\ \|Y - P_V Y\|^2 = \sum_{k=0}^{\infty} |\langle R_Y^k, A^{(i_k)} \rangle|^2$$

特别, 当 $V = R^N$ 时, 有 $Y = P_V Y$, 进而 $\lim_{p \rightarrow \infty} R_Y^p = 0$, 并且

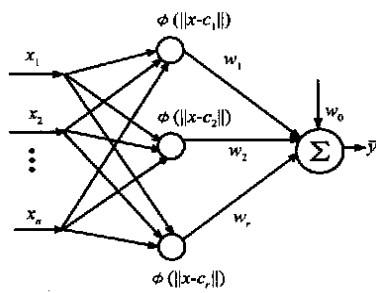


图1 径向基函数网络的结构

$$Y = \sum_{k=0}^1 \langle R_Y^k, A^{(i_k)} \rangle A^{(i_k)} \\ + \|Y - P_V Y\|^2 = \sum_{k=0}^1 |\langle R_Y^k, A^{(i_k)} \rangle|^2$$

虽然自适应投影算法是非线性的, 但是, 分解公式(5)和能量叠加公式(6)却具有正交基分解那样简单的形式.

4 径向基函数网络的设计

设 $\{(x_i, y_i)\}_{i=1}^N$ 为一组训练样本, 其中 $x_i \in R^n, y_i \in R$. 为了从 $\{x_i\}_{i=1}^N$ 中选取 RBF 网络的中心, 我们把 $\langle (+x - x_i) \rangle$ 视为回归因子. 令

$$\begin{aligned} \langle_1 &= (\langle (+x_1 - x_1) \rangle, \langle (+x_2 - x_1) \rangle, \dots, \\ &\quad \langle (+x_N - x_1) \rangle)^T, \quad 1 \leq i \leq N \quad (7) \\ Y &= (y_1, y_2, \dots, y_N)^T \quad (8) \end{aligned}$$

因此, 选取最佳径向基函数中心的问题, 转变成选取 $\{\langle_1, \langle_2, \dots, \langle_N\}$ 的子集, 以使其张成的子空间/最佳地逼近 Y . 为此, 对 Y 在 $\{\langle_1, \langle_2, \dots, \langle_N\}$ 上实施自适应投影算法. 于是, 径向基函数网络的设计过程如下:

(1) 对给定的训练样本集合, 选择径向基函数 \langle_i ;

(2) 用式(7)和式(8)分别计算 \langle_i 和 Y ;

(3) 使用自适应投影算法选取径向基函数的中心, 同时确定了相应的权系数, 具体算法描述如下:

step 1 对 $\{\langle_i\}_{i=1}^N$ 进行能量归一化: $\langle_i = \langle_i / \|\langle_i\|, 1 \leq i \leq N$.

step 2 求 $\langle_{i_0} \in \{\langle_i\}_{i=1}^N$, 使 $|\langle Y, \langle_{i_0} \rangle| = \max_{1 \leq i \leq N} |\langle Y, \langle_i \rangle|$. Y 分解为 $Y = \langle Y, \langle_{i_0} \rangle \langle_{i_0} + R_Y^1$.

step 3 求 $\langle_k \in \{\langle_i\}_{i=1}^N$, 使 $|\langle R_Y^1, \langle_k \rangle| = \max_{1 \leq i \leq N} |\langle R_Y^1, \langle_i \rangle|$.

$$R_Y^1 = \langle R_Y^1, \langle_k \rangle \langle_k + R_Y^{1+}$$

step 4 如果 (a) 残差 $\|R_Y^{k+1}\| + \|Y - \langle Y, \langle_k \rangle \langle_k\| < \epsilon$ 或 (b) 残差衰减率 $\|R_Y^k\| - \|R_Y^{k+1}\| / \|R_Y^k\| + \|Y - \langle Y, \langle_k \rangle \langle_k\| > \epsilon$ 停; 否则, 令 $k = k + 1$, 转 step 3. 其中 ϵ 为一较小的正数.

(4) 删除网络中权系数很小的结点.

设迭代选取的基向量为 $\langle_{i_0}, \langle_{i_1}, \dots, \langle_{i_{L-1}}$ (不妨设它们是互不相同的, 否则, 互相合并), 则选取径向基函数网络的中心分别为 $x_{i_0}, x_{i_1}, \dots, x_{i_{L-1}}$. 即选取径向基函数网络隐层节点的神经元函数分别为 $\langle (+x - x_{i_0}) \rangle, \langle (+x - x_{i_1}) \rangle, \dots, \langle (+x - x_{i_{L-1}}) \rangle$. 这时所对应的网络权系数分别为 $\langle Y, \langle_{i_0} \rangle / \|\langle_{i_0}\|, \langle Y, \langle_{i_1} \rangle / \|\langle_{i_1}\|, \dots, \langle Y, \langle_{i_{L-1}} \rangle / \|\langle_{i_{L-1}}\|$.

这里需要指出的是: 由于自适应投影算法的指数收敛性, 它通常是快速收敛的. 正如 Mallat 在文献[6]中指出的, 分解残量最初是急速下降, 然后出现饱和状态, 因此, 通常在残量快速下降后终止算法.

5 应用实例

下面给出两个 RBF 网络应用于信号处理领域中的实例, 以说明它的实际效果.

(1) 信道均衡器

信道均衡是数字通信中的一个主要问题. 数字通信系统的方框图如图 2 所示, 信道的输入序列 $s(t)$ 是独立的且等概率地取值 1 或 -1, 信道通常用 FIR 滤波器来表示, 其传递函数为

$$c_n(z) = \sum_{i=0}^n c_i z^{-i}$$

其中 c_i 为实数, 且 c_0, c_n 均非零, 加性噪声 $n(t)$ 为高斯白噪声. 均衡器的主要任务是从观测数据

$$Y(t) = (y(t), y(t-1), \dots, y(t-m+1))^T$$

重构输入信号 $s(t-S)$ 的值, m 和 S 分别表示均衡器的阶数和迟延.

传统上, 把均衡问题视为逆滤波问题, 最简单、最常用的均衡器是线性横向均衡器, 然而, 均衡问题是一个固有的非线性问题, 因此, 为了获得最佳的均衡性能, 需要使用非线性滤波器. 由于神经网络的

非线性函数近能力以及形成任意复杂的非线性决策边界的能力, 目前它已成功地应用到均衡问题^[7,8]. 把提出的算法应用于径向基函数网络均衡器的设计, 取得了很好的效果, 通过人们常用的一个例子证实了这点, 这里取 c 为式(2)给出的高斯函数, $R=1$.

取信道的传递函数为

$$c_2 = 0.13482 + 0.18704z^{-1} + 0.13482z^{-2}$$

均衡器的阶数 $m=4$,

迟延 $S=1$. 正如文[8]

所指出的, 这种类型的

信道更接近在实际

通信系统中所遇到的

信道. 图 3 给出了相

应的最佳均衡器、基

于自适应投影学习算

法的 RBF 网络(42 个

RBF 中心) 均衡器和

从训练数据中随机选

取 60 个 RBF 中心的

RBF 网络均衡器^[8]的

性能曲线, 横坐标为

信噪比(SNR), 纵坐标为

误码率(BEA). 对每个

实验用的信噪比, 用

400 个随机选取的训练

样本训练网络, 误码率

是用多于 10^5 个随机选

取的数据点计算的. 由

图可知, 基于自适应投

影学习算法的 RBF 网

络均衡器比基于从训

练数据中随机选取 RBF

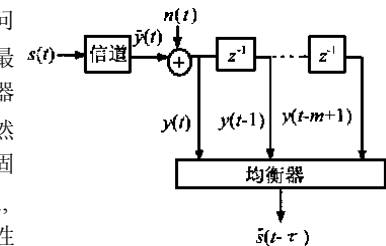


图 2 数据传输系统方框图

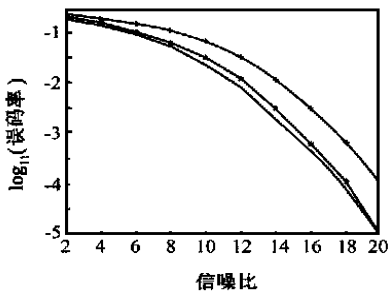


图 3 实线、星号线和加号线分别为最佳均衡器、基于自适应投影学习算法的 RBF 网络均衡器和从训练数据中随机选取 RBF 中心的 RBF 网络均衡器的性能曲线

的单步预报问题. 用 25 个点作为训练数据, 然后用 25 个点作为检验数据, 取径向基函数为高斯函数 $\phi(x) = e^{-4x^2}$, 在径向基函数网络的设计过程中, 取迭代终止准则为 (a) (见 step 4), 令 $E = 10^{-3}$. 在图 4 中, (a) 和 (b) 分别给出了基于自适应投影学习算法的 RBF 网络(8 个 RBF 中心)和基于从训练数据中随机选取 12 个 RBF 中心的 RBF 网络^[8]的预报曲线, 其中虚线为最终预报结果, 而实线为实际结果. 由图可见, 基于自适应投影学习算法的 RBF 网络比基于从训练数据中随机选取 RBF 中心的 RBF 网络用于时间序列的预报, 具有更高的预报精度.

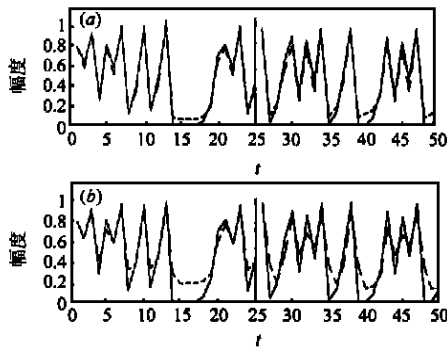


图 4 (a) 基于自适应投影学习算法 RBF 网络的预报曲线. (b) 基于随机选取 RBF 中心的 RBF 网络的预报曲线.

6 结论

构造 RBF 网络的最大困难是径向基函数的数目和中心位置的合理选取. 本文提出的构造和训练径向基函数网络的自适应投影算法, 它建立在自适应投影原理的基础上, 充分利用训练数据, 迭代地确定径向基函数的数目、中心的位置和网络的权系数, 所生成的径向基函数网络具有体积小、收敛快和数值稳定等优点, 这些通过两个不同的信号处理领域应用的例子得到了进一步证实.

作者简介:



张苗生 副教授, 在职博士生, 1989 年获西安交通大学应用数学专业硕士学位. 主要从事神经网络、小波分析、联合时频分析及非平稳信号处理等领域的研究工作.



刘贵忠 教授, 博士生导师, 1989 年获荷兰 Eindhoven 大学博士学位. 主要从事非平稳信号处理、多媒体技术、通信及测控等领域的研究工作.

(下转第 134 页)

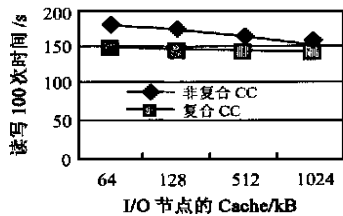


图 5 两用户随机共享读写

Cache 机制下, 系统充分使用了网络连接的局部内存, 使得系统读写性能明显优于没有 Cooperative Cache 机制时的系统性能。由于 Cooperative Cache 机制充分利用了网络共享内存, 所以数据缓冲区基本是够用的, 故单机 Cache 的增大并不会对系统性能有所改变, 系统性能较为稳定。而无 Cooperative Cache 机制的共享访问时间则随 Cache 的增大逐渐减少, 显示出 I/O 节点靠本地缓冲难以保证数据的充分缓冲, 所以当单机 Cache 增大到能满足系统缓冲要求时, 系统性能较好, 但若 Cache 不足时, 系统性能就较差。由于单机内存有限, 若无法充分利用它机内存, 将难以应付大量数据缓冲的要求。这显示出复合 Cooperative Cache 机制能有效提高系统资源利用率, 提高系统效率。

5 结束语

本文介绍了基于 NOW 的并行文件系统 PARFSNOW++ 的设计中涉及 Cooperative Cache 技术。经过测试表明, 这些技术的引入确实提高了并行文件系统的效率。

参考文献:

- [1] T. E. Anderson, M. D. Culler, D. A. Patterson et al. A case for NOW (Networks of Workstations) [J]. IEEE Micro, February 1995: 54- 64.
- [2] M. Nelson, B. Welch and, J. Ousterhout. Caching in the sprite network file system [J]. ACM Trans. on Computer Systems, February 1988, 6 (1): 45- 55.
- [3] D. Kotz and, N. Nieuwejaar. Flexibility and Performance of Parallel File Systems [J]. ACM Operating Systems Review, April 1996, 30(2): 63 - 73.
- [4] 李群, 谢立, 孙钟秀. 并行文件系统的设计 [J]. 计算机科学, 23 (4): 36- 39.
- [5] Erik Hagersten et al. DDM2A Cache2Only Memory Architecture [J]. IEEE Computer Sep. 1992: 44- 54.
- [6] A. Leff, J. L. Wolf, S. Yu. Efficient LRU2 Based Buffering in a LAN Re2 mote Caching Architecture [J]. Transactions on Parallel and Distributed Systems, February 1996, 7(2): 191- 206.

作者简介:

赵 欣 1974 年出生。1996 年获南京大学计算机系学士学位, 1998 获南京大学计算机系硕士学位。现在南京大学计算机系攻读博士学位。研究方向: 网络与分布式系统。发表文章 10 余篇。

陈道蓄 1982 年毕业于南京大学。现任南京大学计算机系教授, 系主任。主要研究方向: 分布式处理和并行计算。在国内外发表论文 30 余篇, 合作出版著作 2 本, 并 4 次获教育部和江苏省科技进步奖。

谢 立 1964 年毕业于南京大学。现任南京大学计算机系教授, 南京大学副校长。主要研究方向: 分布式处理和并行计算。在国内外发表论文 100 余篇。

(上接第 122 页)

参考文献:

- [1] D. Hush, B. Home. Progress in supervised neural networks [J]. IEEE Signal Processing Magazine, January 1993: 8- 39.
- [2] J. Park, I. W. Sandberg. Approximation and radial basis function networks [J]. Neural Computa., 1993, 5(1): 305- 316.
- [3] J. Park, I. sandberg. Universal approximation using radial basis function networks [J]. Neural Computa., 1991, 3(2): 246- 257.
- [4] N. B. Karayiannis, G. W. Mi. Growing radial basis neural networks: merging supervised and unsupervised learning with network growth techniques [J]. IEEE Trans. Neural Networks, 1998, 8(6): 1492 -

1506.

- [5] S. Chen, C. F. N. Cowan, P. M. Grant. Orthogonal least squares learning algorithm for radial basis function networks [J]. IEEE Trans. Neural Networks, 1991, 12(2): 302- 309.
- [6] S. Mallat, Z. Zhong. Matching pursuits with time2frequency dictionaries [J]. IEEE Trans. Signal Processing, 1993, 41(12): 3397- 3415.
- [7] R. Parisi, E. D. D. Claudis, G. Orlandi, B. D. Rao. Fast adaptive digital equalization by recurrent neural networks [J]. IEEE Trans. Signal Processing, 1997, 45(11): 2731- 2739.
- [8] S. Chen, G. J. Gibson, C. F. N. Cowan, P. M. Grant. Reconstruction of binary signal using an adaptive radial2basis2function equalizer [J]. Signal Processing, 1991, 22(1): 77- 93.