

文章编号:1000-582X(2003)08-0048-04

Legendre 小波神经网络*

郑小洋, 叶仲泉

(重庆大学 数理学院, 重庆 400044)

摘要:在 BP 神经网络的基础上, 结合 Legendre 小波构造了 Legendre 小波神经网络。由于 Legendre 小波在区间 $[0, 1]$ 上具有分段表达式并且为多项式的特点, 因而构造的 Legendre 小波神经网络有结构简单、收敛速度快等优点。以神经网络的 BP 算法作为 Legendre 小波神经网络的学习算法, 用有 6 个 Legendre 小波基函数的 Legendre 小波神经网络对一个函数进行逼近分析, 得到了较好的逼近效果。

关键词:Legendre 小波; 神经网络; BP 算法; Legendre 小波神经网络

中图分类号:TP183; O174.2

文献标识码:A

神经网络以其快速的并行处理能力和其强有力的学习能力而获得越来越广泛的重视, 神经网络系统最主要的特征是大规模模拟并行处理、信息的分布式存储, 高度的容错性和自组织、自学习及实时处理, 它可以直接输入样本, 信息处理分布于大量神经元的互连之中, 并且具有冗余性。虽然神经网络具有很多优点, 但也存在一些缺点。从大量的实际应用来看, 收敛速度慢, 学习时间长, 易陷入局部极小, 学习过程中产生大的振荡, 甚至达不到收敛精度。而小波神经网络是基于小波分析所构造的一种新型的神经网络模型^[1], 它结合了小波变换良好的时频局域化性质及神经网络的自学习功能, 因而具有较强的逼近、容错能力。小波神经网络相比于前向的神经网络, 它有明显的优点^[2]: 首先小波神经网络的基元和整个结构是依据小波分析理论确定的, 可以避免 BP 神经网络等结构设计上的盲目性; 其次小波神经网络有更强的学习能力, 精度更高; 最后对同样的学习任务, 小波神经网络结构更简单, 收敛速度更快。在 BP 神经网络的基础上结合 Legendre 小波构造了 Legendre 小波神经网络。

1 神经网络

已经提出了多种神经元模型, 如神经元数理模型, 分层神经网络模型, Hopfield 模型, 认知模型等。从本质上说神经网络是一种巨型非线性动力系统, 功能强大的计算产生于它的潜在动力学演化过程, 在神经网络与外界环境的相互作用过程中, 神经网络逐步调整自己以适应环境, 这种调整源于神经网络的自学能力。

在构造的 Legendre 小波神经网络的学习算法上用了 BP 算法, 故笔者在此主要介绍 BP 算法和训练网络的学习阶段。BP 算法的主要思路是: 若能求出误差 E 对各个神经元输出的偏导数, 就可以算出 E 对所有连接权的偏导数, 因而就可以利用梯度下降法来修改各个连接权。在训练网络的学习阶段^[3], 给定模式作为网络的输入, 要求网络通过调节所有的连接权系数和各神经元的阈值, 使得在输出层神经元上得到需要的理想输出。一旦网络完成这种调节, 则再给出另一个模式和相应的理想输出, 要求网络继续完成对这对模式的学习。实际上就是要求网络寻找单一的一组连接权系数和偏置, 以满足给定的所有输入输出模式对。

2 Legendre 小波神经网络

已知, 前向多层网络是基函数神经网络, 一般采用极性函数作为基函数。而小波可提供区别函数的多分辨近似, 提供空间和频率的局部化, 以小波分析理论为基础构造的小波神经网络更适合于学习局部非线性和快速变化的函数。

2.1 小波分析基本知识

如果 $\psi \in L^2(\mathbb{R})$ 满足“容许性”条件:

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\psi(\omega)|}{|\omega|} d\omega < +\infty \quad (1)$$

那么称 ψ 是一个“基小波”或“母小波”, 这里“容许性”条件一般把 ψ 限制为一振荡衰减且具有紧支集的窗函数。小波 $\psi(x)$ 通常由其伴随即尺度函数 $\varphi(x)$ 生成:

* 收稿日期: 2003-02-04

作者简介: 郑小洋 (1972-), 男, 重庆人, 重庆大学硕士研究生, 主要从事大系统优化理论及应用、小波分析和神经网络的研究。

$$\begin{aligned} \varphi(x) &= \sqrt{2} \sum_n h_n \varphi(2x - n) \\ \psi(x) &= \sqrt{2} \sum_n g_n \varphi(2x - n) \end{aligned} \quad (2)$$

由 ψ 生成的函数族 $\{\psi_{a,b}\}$ 为:

$$\psi_{a,b}(x) = |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \quad (3)$$

其中 $a, b \in R$ 且 $a \neq 0$, 称为连续小波。设 $f(x) \in L^2(R)$, 则称

$$W_T(a, b) = |a|^{-\frac{1}{2}} \int_{-\infty}^{+\infty} f(x) \overline{\psi\left(\frac{x-b}{a}\right)} dx \quad (4)$$

为小波变换。当 $a = 2^{-j}, b = 2^{-j}k (j, k \in Z)$ 时有离散小波, 对于离散小波可以通过同一函数 ψ 的平移和伸缩而构成的 $L^2(R)$ 中的正交基, 它满足如下条件^[4]:

1) 对于一切的整数 k , 函数 $\varphi(x - k)$ 相互正交。

2) 函数 φ 是一尺度函数, 对于任意的整数 $j, \{\varphi(2^j x - k)\}$, 构成了它线性张成的子空间的一组正交基。

3) 小波函数定义为 $\psi, \{\psi(2^j x - k)\}$, 构成了 $L^2(R)$ 中的一组正交基。能够证明 $\{\varphi(2^{j_0} x - k), \psi(2^j x - k) j \geq j_0\}$ 也能构成 $L^2(R)$ 中的一组正交基。小波子空间定义如下:

$$W_j = \{\psi(2^j x - k), k \in Z\} \quad (5)$$

并满足:

$$\begin{aligned} V_j &= V_{j-1} \oplus W_{j-1} \\ W_j \cap W_i &= \{\phi\}, j \neq i \end{aligned} \quad (6)$$

这里符号“ \oplus ”表示直和。

由多分辨分析的知识知道, 任意小波产生一个 $L^2(R)$ 中的直和分解, 对于每一 $j \in Z$, 我们考虑以下 $L^2(R)$ 中的闭子空间:

$$V_j = \dots \oplus W_{j-2} \oplus W_{j-1} \quad (7)$$

因此一维函数 $f(x)$ 的小波级数表示如下:

$$\begin{aligned} f(x) &= \sum_{k \in Z} a_{j_0 k} \varphi_{j_0 k}(x) + \sum_{j > j_0} \sum_{k \in Z} b_{j k} \psi_{j k}(x) \\ \varphi_{j_0 k}(x) &= 2^{j_0/2} \varphi(2^{j_0} x - k), \\ \psi_{j k}(x) &= 2^{j/2} \psi(2^j x - k) \end{aligned} \quad (8)$$

小波系数 $a_{j_0 k}$ 和 $b_{j k}$ 如下:

$$\begin{aligned} a_{j_0 k} &= \langle f(x), \varphi_{j_0 k}(x) \rangle \\ b_{j k} &= \langle f(x), \psi_{j k}(x) \rangle \end{aligned} \quad (9)$$

这里符号“ $\langle \cdot \rangle$ ”表示内积。

由此可知小波级数把函数 $f(x)$ 分解成不同频率信号的组合。

2.2 Legendre 小波

这里对 Legendre 小波作介绍^[5], Legendre 小波 $\psi_{nm}(t) = \psi(k, \hat{n}, m, t), k = 2, 3, \dots, \hat{n} = 2n - 1, n = 1, 2, 3, \dots, 2^{k-1}, m$ 是 Legendre 多项式的阶, t 为时间。它们定义在区间 $[0, 1)$ 上, 满足下式:

$$\psi_{nm}(t) = \begin{cases} \left(m + \frac{1}{2}\right)^{1/2} 2^{k/2} L_m(2^k t - \hat{n}), & \frac{\hat{n}-1}{2^k} \leq t < \frac{\hat{n}+1}{2^k} \\ 0, & \text{其它} \end{cases} \quad (10)$$

$L_m(t)$ 就是著名的 m 阶 Legendre 多项式, 满足下面的递推公式:

$$\begin{aligned} L_0(t) &= 1 & L_1(t) &= t \\ L_{m+1}(t) &= \frac{2m+1}{m+1} t L_m(t) - \frac{m}{m+1} L_{m-1}(t), \\ & & m &= 1, 2, 3, \dots \end{aligned} \quad (11)$$

对不相同的 n 的 2 个 Legendre 小波函数显然正交。当 n 相同而 m 不相同, 因为 Legendre 多项式是 $[-1, 1]$ 上的正交多项式, 所以有:

$$\int_0^1 \psi_{nm_1}(t) \psi_{nm_2}(t) dt = \int_{\frac{\hat{n}-1}{2^k}}^{\frac{\hat{n}+1}{2^k}} \left(m_1 + \frac{1}{2}\right)^{1/2} 2^{k/2} \cdot$$

$$L_{m_1}(2^k t - \hat{n}) \left(m_2 + \frac{1}{2}\right)^{1/2} 2^{k/2} L_{m_2}(2^k t - \hat{n}) dt =$$

$$\left(m_1 + \frac{1}{2}\right)^{1/2} \left(m_2 + \frac{1}{2}\right)^{1/2} 2^k \int_{-1}^1 L_{m_1}(u) \cdot$$

$$L_{m_2}(u) du = 0, u = 2^k t - \hat{n}, m_1 \neq m_2$$

得 Legendre 小波集是一正交集。

2.3 Legendre 小波神经网络

由 Legendre 小波可知, 一个定义在 $[0, 1)$ 上的函数 $f(t)$ 可展开成如下式:

$$\begin{aligned} f(t) &= \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} c_{nm} \psi_{nm}(t) \\ c_{nm} &= \langle f(t), \psi_{nm}(t) \rangle \end{aligned} \quad (12)$$

如果作近似逼近, 则可得下式:

$$f(t) \approx \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} \psi_{nm}(t) = C^T \psi(t) \quad (13)$$

这里 C 和 $\psi(t)$ 是 $2^{k-1} M \times 1$ 矩阵:

$$\begin{aligned} C &= [c_{10}, c_{11}, \dots, c_{1M-1}, c_{20}, \dots, c_{2M-1}, \dots, c_{2^{k-1}0}, \dots, c_{2^{k-1}M-1}]^T \\ \psi(t) &= [\psi_{10}(t), \psi_{11}(t), \dots, \psi_{1M-1}(t), \psi_{20}(t), \dots, \\ & \quad \psi_{2M-1}(t), \dots, \psi_{2^{k-1}0}(t), \dots, \psi_{2^{k-1}M-1}(t)] \end{aligned}$$

用 Legendre 小波基函数作为神经网络的激励函数, 由式(13)可构造出 Legendre 小波神经网络, 其网络结构如图 1。

网络结构:

- 1) 输入层: 输入数字化的原始信号。
- 2) 预处理层: 对输入的数字化的原始信号分为 2^{k-1} 组, 这些分组后的数字信号分别进入相应的 Legendre 小波基函数进行训练。预处理层的输出等于其相应的输入。

3) 隐含层: 类似于 BP 神经网络的隐含层, 它分为 2^{k-1} 组结点, 每一组有 M 个 Legendre 小波基函数, 分别接受预处理层处理后的数字信号。隐含层结点的权值就是 Legendre 小波系数的逼近。

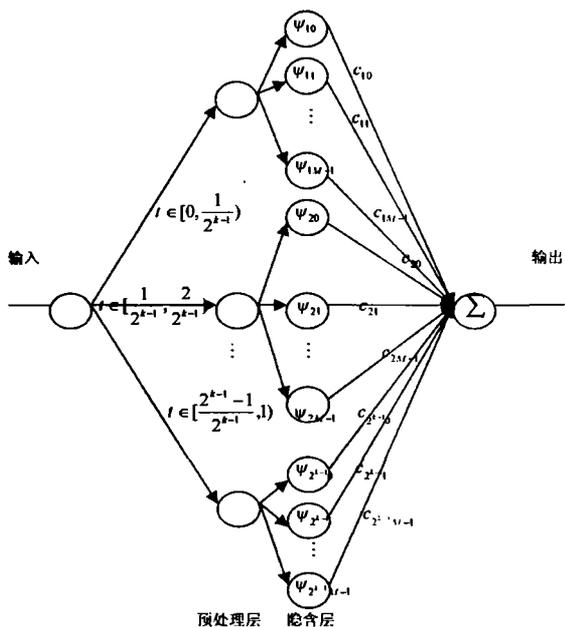


图1 Legendre小波神经网络结构图

4) 输出层:类似于BP网络中的输出层,接受隐含层的输出,输出层是线性结点,进行求和运算,输出结果。

现在给出具体的训练算法:

1) 网络的训练:网络的训练采用BP算法,对于定义在区间 $[0,1)$ 上的函数 $f(t)$,由式(13),记为:

$$g(t) = \sum_{n=1}^{2^{k-1}} \sum_{m=0}^{M-1} c_{nm} \psi_{nm}(t) = C^T \psi(t) \quad (14)$$

则通过适当选取权值 c_{nm} ,上式可作为 $f(t)$ 的逼近。若已知训练数据集,权值可通过最小化均方误差来得到^[6]:

$$(\hat{c}_{10}, \dots, \hat{c}_{1M-1}, \hat{c}_{20}, \dots, \hat{c}_{2M-1}, \dots, \hat{c}_{2^{k-1}0}, \dots, \hat{c}_{2^{k-1}M-1}) = \arg \min_{c_{10}, \dots, c_{1M-1}, c_{20}, \dots, c_{2M-1}, \dots, c_{2^{k-1}0}, \dots, c_{2^{k-1}M-1}} \varepsilon_N(f, g) \quad (15)$$

式(15)中的

$$\varepsilon_N(f, g) = \frac{1}{N} \sum_{i=1}^N (f(t_i) - g(t_i))^2 \quad (16)$$

$$\frac{\partial \varepsilon_N(f, g)}{\partial c_{nm}} = 0 \quad (17)$$

式(15)可由式(17)求得。

2) 初始化:取小的随机数初始化权系数。

3) 循环:设定迭代次数,读入数据进行网络训练,一旦网络训练好就得到了所需要的权系数。

4) 保留 Legendre 小波神经网络的权系数值,结束训练。

由于 Legendre 小波基函数在区间 $[0,1)$ 上具有分段表达式并且为多项式的特点,因而构造的 Legendre 小波神经网络结构简单、收敛速度快。对以上由 Legendre 小波基函数构成的小波神经网络训练时可分 2^{k-1} 组进行,每一组有 M 个 Legendre 小波基函数,从而在训练小波神经网络上可以减小训练规模,而训练后

的最后结果又能体现整个 Legendre 小波神经网络的特性。如果要提高精度,可增加 M 或 k 的值,或者同时增加 M 和 k 的值。 k 值的增加就是进一步细分区间 $[0, 1)$, M 值的增加就是增加相应细分区间上的多项式的最高阶的次数。

当 $M = 3$ 和 $k = 2$ 时,有6个 Legendre 小波基函数如下:

$$\begin{aligned} \psi_{10} &= 2^{1/2}, \\ \psi_{11} &= 6^{1/2}(4t - 1), 0 \leq t < \frac{1}{2} \end{aligned} \quad (18)$$

$$\psi_{12} = (10)^{1/2} \left[\frac{3}{2}(4t - 1)^2 - \frac{1}{2} \right]$$

$$\begin{aligned} \psi_{20} &= 2^{1/2}, \\ \psi_{21} &= 6^{1/2}(4t - 3), \frac{1}{2} \leq t < 1 \end{aligned} \quad (19)$$

$$\psi_{22} = (10)^{1/2} \left[\frac{3}{2}(4t - 3)^2 - \frac{1}{2} \right]$$

用这6个 Legendre 小波基函数作为神经网络的激励函数,就可构造出一个有6个 Legendre 小波基函数的 Legendre 小波神经网络,其网络结构如图2。

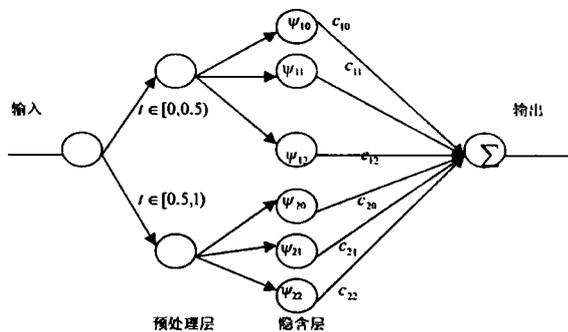


图2 有6个基函数的 Legendre小波神经网络结构图

这时有6个权值需要确定。

3 应用

用上面构造的有6个 Legendre 小波基函数的 Legendre 小波神经网络对一函数进行逼近分析,此函数为:

$$f(x) = e^{-4x} \cos(12x) + \cos(4x), x \in [0,1) \quad (20)$$

给定一训练数据集:

$$T_{N_0} = \{x_i, f(x_i)\}_{i=1}^{N_0} \quad (21)$$

以上采用的训练数据是均匀的,即均匀的在函数定义域上取 N_0 个点。Legendre 小波神经网络的具体的训练算法按上面介绍的来进行训练。为评价逼近效果,采用均方误差作为评价指标:

$$MSE = \frac{1}{M} \sqrt{\sum_{i=1}^M (g(x_i) - f(x_i))^2} \quad (22)$$

式(22)中的 M 为测试数据集个数。

当训练数据 $N_0 = 200$, 迭代次数 N 为 20 次时有: 图 3, 图 4 中的实线表示原函数, 虚线表示逼近结果。

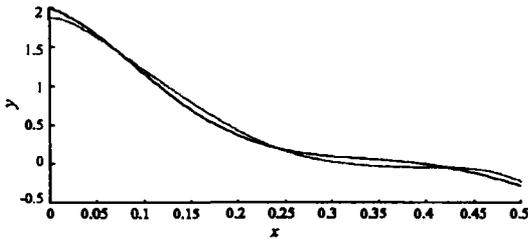


图 3 在区间 $[0, 0.5)$ 上原函数和逼近函数的图形

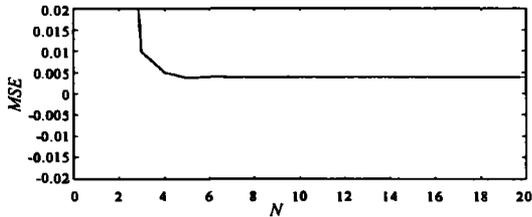


图 4 在区间 $[0.5, 1)$ 上原函数和逼近函数的图形
均方误差 MSE 随迭代次数 N 变化的图形如图 5, 图 6。

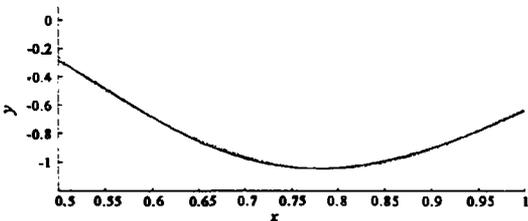


图 5 在区间 $[0.5, 1)$ 上均方误差随迭代次数 N 变化的图形

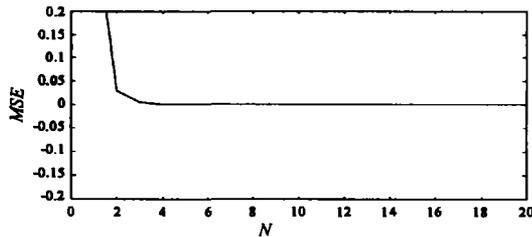


图 6 在区间 $[0.5, 1)$ 上均方误差随迭代次数 N 变化的图形

由图 5、图 6 可知经过大约 6 次迭代后均方误差基本不再发生变化, 这说明收敛速度快。

经 20 次迭代训练后, 在区间 $[0, 0.5)$ 和区间 $[0.5, 1)$ 上的均方误差分别为:

0.0040; $8.5625e - 005$

用 50 个测试数据对训练好后的 Legendre 小波神经网络进行评价, 在区间 $[0, 0.5)$ 和区间 $[0.5, 1)$ 上的均方误差分别为:

0.0068; $1.1848e - 004$

权值为: $[0.34417 \quad -0.42875 \quad 0.10317 \quad -0.5818 \quad -0.076321 \quad 0.12042]$, 这 6 个权值就是对 6 个小波系数的一个逼近。

4 结 论

用 Legendre 小波构造了 Legendre 小波神经网络, 并用有 6 个 Legendre 小波基函数的一个 Legendre 小波神经网络对一函数进行了逼近分析, 得到了较好的效果。现有的小波神经网络的学习算法以梯度下降类为主, 而梯度算法有收敛速度慢等缺点, 因此可以对小波神经网络的学习算法作进一步的研究分析, 从而提高小波神经网络的性能。

参考文献:

- [1] 蔡文生, 任琴, 邵学广, 等. 小波神经网络用于色谱数据的滤噪与压缩[J]. 分析试验室, 1998, 17(6): 1-4.
- [2] 刘建春, 王正欧. 一种小波神经网络的快速学习算法及其应用[J]. 天津大学学报, 2001, 34(4): 455-458.
- [3] 黄国建, 王建华, 芦国强. 提高人工神经网络 BP 算法收敛速度的几种方法[J]. 华东船舶工业学院学报, 1994, 8(2): 23-29.
- [4] LIU G P, BILLINGS S A, ADIRKAMANTHAN V K. Nonlinear system identification using Wavelet Networks[J]. International Journal of Systems Science, 2000, 31(12): 1531-1541.
- [5] AZZAGHI M R, YOUSEFI S Y. The Legendre Wavelets Operational Matrix of Integration[J]. International Journal of Systems Science, 2001, 32(4): 495-502.
- [6] 顾成奎, 王正欧. 基于多分辨分析神经网络的函数逼近[J]. 天津大学学报, 2001, 34(1): 119-123.

Legendre Wavelets Neural Network

ZHENG Xiao-yang, YE Zhong-quan

(College of Mathematics Sciences, Chongqing University, Chongqing 400044, China)

Abstract: Based on BP neural network, a Legendre wavelets neural network is constructed with Legendre wavelets. The Legendre wavelets defined on the interval $[0, 1)$ has the features of piecewise expression and being polynomials, so the Legendre wavelets neural network has the advantage of simple structure and fast convergence speed. Trained by BP algorithm, the Legendre wavelets neural network with six wavelet basis functions is used to learn a function and better approximate result is obtained.

Key words: legendre wavelets; neural network; BP algorithm; legendre wavelets neural network

(编辑 吕赛英)