33-38
第16卷第1期 红外与毫米波学报 Vol. 16, No. 1
1997年2月 J. Infrared Millim. Waves February, 1997

用于信号或函数逼近的子波神经网络快速收敛初值设定法*

宋湘黔 陈建铭 戚飞虎
(上海交通大学光纤技术研究所,上海,200052)

TN919 TN911.1 TN911.7

A 摘要 针对目前用于信号(函数)逼近的子波神经网络的收敛速度慢的问题,提出一种子波网络的初值设定方法,使子波网络的收敛速度大大提高,通过对线性多项式、指数、三角函数及某些分段函数的实验以及与不同网络的对比,表明此方法具有很强的普适性.

关键词 子波,子波变换,神经网络,函数逼近. 信号处理

⑦

引言

信号(函数)的逼近对于信号的传输、压缩具有重要意义.由于神经网络的人工智能的特性,使它在对非线性映射的逼近问题上显示了优势,比较成熟的BP网络多基于Sigmoid函数,从对函数的表示角度看,它是一类次优网络[1],随着子波理论的兴起,子波神经网络成为研究的热点,Hurold H. Szu 构造了子波神经网络模型,并用它对信号进行逼近和分类[2]. Q. Zhang 等采用类似结构进行了函数逼近[3],证实了子波神经网络的可行性,但是收敛速度不能令人满意,Q. Zhang 及国内同仁的实验[4]中,均需上万次的迭代来实现对一维信号(函数)的逼近.实验表明,网络初值的设定对收敛速度有很大的影响.在上述实验中,利用信号函数与子波函数的相关结果如卷积等,设定子波函数的伸缩及平移系数,网络连结权值的初值则随意或设零.我们的实验表明,子波神经网络有较强的自适应能力,它能在较大范围内调整子波函数系的伸缩因子,获得相应于该信号的一组最佳基,因此,我们提出一种注重网络连结权值的初值设定方法,使网络的收敛速度大大提高,对相同信号(函数)的逼近,只需千次迭代,大大提高了算法的实用性.

1 子波及子波变换

子波分析的基本思想是用一族函数表示或逼近一信号或函数,这一族函数称子波函数系,它是通过一基本子波函数的不同尺度的平移和伸缩构成.记基本波函数为 $h(x)$,伸缩和平移因子分别为 a 和 b,则子波变换基底为

$$n_{a,b}(x) = |a|^2 h\left(\frac{x-b}{a}\right), \tag{1}$$

函数 $f(x)$ 连续子波变换定义为 L^2 的希尔伯特空间上的内积,即

* 国家攀登资助项目
本文1996年1月26日收到,修改稿1996年5月7日收到

$$Wf(a,b)=\int_{-\infty}^{+\infty}h_{a,b}(x)f(x)d_x=<h_{a,b}(x),f(x)>. \tag{2}$$

它对应于 $f(x)$在函数族 $h_{a,b}(x)$上的分解，其中母波函数 $h(x)$的傅里叶变换应满足容许条件

$$c_h=\int_{-\infty}^{+\infty}\frac{|h(w)|^2}{|w|}d_w<+\infty>\text{或}\int_{-\infty}^{+\infty}h(x)d_x=0\ . \tag{3}$$

由式(2)定义的子波变换可通过对其伸缩因子 a 和平移因子 b 的采样离散化而得到离散子波变换.

子波变换的突出优点是时宽带宽乘积很小，且在时间和频率轴上都很集中，具有放大、缩小、平移等功能，使它的信号高频时使用短窗口，低频时采用宽窗口，体现了多分辨率的思想，这对于突变信号、非平稳信号的分析是非常重要的. 它可将任一信号分解成时间和频率的独立贡献，同时又不失原有信号所包含的信息. 正是由于这些突破传统傅里叶分析的特性，使得它在信号的表示或逼近上的应用显示了很大的优势.

2 用于信号表示或逼近的子波神经网络

我们采用图1所示的子波神经网络，若是对信号逼近，则图中输入为 t，输出为 $f(t)$. 引入 $\bar{f}$ 为 $f(x)$的均值估计. 则得到子波连续神经网络的响应方程

$$\hat{f}(x)=\sum_{i=l}^{N}W_i h_{a_i b_i}(\frac{x-b}{a_i}+\bar{f}). \tag{4}$$

需要学习的参数有 W_i、b_i、a_i 参数的学习过程可视为：子波变换的投影能量自适应地向 N 时频分布点$[a_i,b_i]$，$i=1,2...,N$ 进行集中的过程. 网络参数 W_i、b_i、a_i 的迭代过程可通过能量函数最小化实现. 我们采用 LMS 能量函数

$$E=\frac{1}{2}\sum_{i=l}^{P}[f(x)-\hat{f}(x)]^2. \tag{5}$$

其中 P 为采样的样点个数，$f(x)$为对应于 x 的函数的目标输出. 采用共轭梯度最小化，即

$$w_i^{k+1}=w_i^k-\eta_1\,\partial E/\partial w_i, \tag{6}$$

$$a_i^{k+1}=a_i^k-\eta_2\,\partial E/\partial \alpha_i, \tag{7}$$

$$b_i^{k+1}=b_i^{k+1}-\eta\partial E/\partial b_i. \tag{8}$$

我们采用子波基函数 $h(x)=\cos(1.75x)\exp(-s^2/2)$，可以证明它满足容许条件(3)，令 $x'_j=(x_j-b_i)/a_i$，则迭代方程(6)、(7)、(8)中的梯度为

$$\frac{\partial E}{\partial w_i}=-\sum_{j=l}^{P}[f(x_j)-\hat{f}(x_j)]\cos(1.75x'_j)\exp(x'^2_j)/2, \tag{9}$$

$$\begin{aligned}\frac{\partial E}{\partial \alpha_i}=-\sum_{j=l}^{P}[f(x_j)-\hat{f}(x_j)]w_i[&1.75\sin(1.75x'_j)\exp(x'^2_j)/2\ x'/a_j\\&+\cos(1.75x'_j)\exp(x'^2_j)/2\ x'^2_j/\alpha_j),\end{aligned} \tag{10}$$

$$\frac{\partial E}{\partial b_i}=-\sum_{j=1}^{P}[f(x_j)-\hat{f}(x_j)]w_i[1.75\sin(1.75x'_j)\exp(x'^2_j)/2\ x'/a_i$$
$$+\cos(1.75x'_j)\exp(x'^2_j)/2\ x'^2_j/a_j). \tag{11}$$

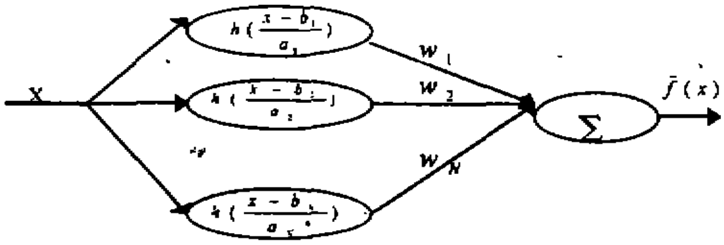

图 1　连续子波神经网络

Fig. 1 Continuous wavelet neural network

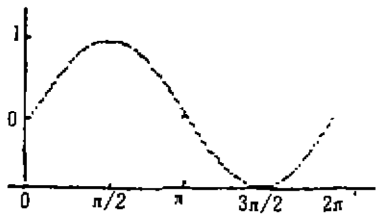

图 2　函数 sin(x)300 次迭代(err=0.0654)

Fig. 2 Approximation of sin(x) (300iterations, err=0.0654)

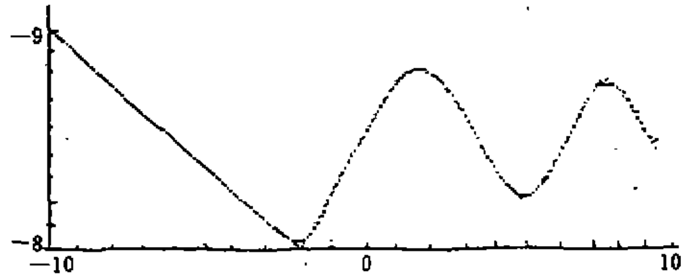

图 3　函数(12)1500 次迭代(err=0.0392)

Fig. 3 Approximation of function(12) (1500 iterations, err=0.0392)

我们得出子波网络初值设定方法为:(1)若网络采用 N 个子波,则我们把 P 个取样点分成 N 个取样区间,前 N-1 段,采样点为不大于(P/N)的最大,整数个记为 M. 第 N 段中取样点为[P-M(M-1)],记为 M_N 个. (2)对每个区间,计算函数(信号)的平均值. $f_i=\sum_{i=1}^{M_i}f(x_i)/M_i$,对 $i=1,\ldots,N$-1,M;对 $i=N$,$M_i=M_N$,将 W 的初值定为 $w_i=f_i$. (3)b 的初值均匀分布于取样区间,即 $b_i=(xN-x_1)/N\ i$;a 的初值则取 2～10 之间的某个值. 可见,我们注重于 W 的初值设定.

以上的初值设定法以及它保证波网络的快速收敛性,主要是依据于以下的理论分析和实验尝试。

(1) 子波函数 $h(x)=\cos(1.75x)\exp(-x2/2)$有很好的局域性. 在 $x=0$ 时函数值达到最大值 1,随着 X 的增大或减小,函数值迅速向零减小. 对于不同的平移和伸缩系数 a 和 b 而生成的一族子波函数,可以近似认为相互独立,即对于需要逼近的信号或函数,投影基底近似独立正交. 而 Sigmoid 函数平缓地分布于整个空间,一族平移的 Sigmoid 函数不能构成一组独立正交基.

(2) W 相当于投影基底的系数,对于非正交基底,不同的 W 的调整互相之间有很大的影响,从而造成网络的收敛速度比正交基底的网络慢.我们对比了子波神经网络和典型的一般 BP 网络、冲量算法[5]BP 网络,实验结果见后,由于子波变换的优越性和上述原因,子波神经网络应具有比常用的 RBF 网络、样条函数网络更快的收敛速度.由于篇幅有限,本文未给出比较结果.

(3) 对于同是子波网络,若 W 随意设定初值或零,则对于需要逼近的信号或函数,取样空间的 N 各子波函数需要根据各处的函数幅度调节 W,调节幅度可能很大,并且各处的子波函数互相影响,所以,需要多次的协调才能达到最佳值.而上述初值的设定法,将初始子波幅度选为某段区间的函数平均值,减少了 W 的调节幅度,使网络能在较少的的步骤内收敛到最佳值.

(4) 我们的实验表明,子波神经网络有较强的自适应能力,它能在较大范围内调整子波函数的伸缩因子,获得相应于该信号的一组最佳基[6],故对不同样的基底数,子波网络可以达到比其它函数基底网络更小的误差,也使我们简化 a 值的设定,根据实验尝试,取 2～10 之间的某个值,根据均匀标尺的思想,我们将子波基均匀分布在取样空间,即得以上 b 的初值设定.

3 实验结果及分析

我们对多种函数进行了实验,均取得了满意的结果.为了比较,我们对 W. Zhang[3]文章中的函数进行逼近.这一函数为

$$f(x)=\begin{cases}-2.18x-12.864, & -10\le x<-2\\ 4.246x, & -2\le x<0\\ 10e^{-0.05x-0.5}sin[(0.03x+0.7)x]. & 0\le x\le 10\end{cases} \tag{12}$$

取样点为[－10＋10]上均匀分布的 200 个点,取同样的误差函数

$$\mathrm{err}=\sqrt{\sum_{j=l}^{P}[\hat{f}(x_j)-f(x_j)]^2/\sum_{j=l}^{P}[f(x_i-\bar{f})]^2},$$

其中 $\bar{f}=\frac{1}{p}\sum_{j=l}^{P}f(x_j)$,取 $N=10$,a 的初值为 2.0. $\eta_1=0.01$,$\eta_2=0.104$,$\eta_3=0.104$. 经 1500 次迭代,err＜0.04,而 Q. zhang 的实验中,经 10000 次迭代,err＝0.05057,在 1000 次迭代时,a 的值如表 1,最大值为最小值的 2.8 倍,可见子波神经网络有很强的自适应调节能力,使 a 达到最佳,也证明我们初值假定中对 a 的简单设定的合理性.表 2 显示了对一些函数的实验结果,取样点为[0,2π]上均匀的 100 个点,$N=10$. 表 3 是对同样函数,用一般 BP 网络和冲量算法 BP 网络的实验结果,采用单隐层结构,隐单元节点为 8 个.图 2、3 显示了一些函数逼近的结果,可见这种方法的普适性.

表 1　子波网络的伸缩系数

Table 1 Dilation values of wavelet network

i	1	2	3	4	5	6	7	8	9	10
a_i	3.7734	2.5834	2.0320	1.9250	2.9214	1.3067	2.4606	2.1375	1.5874	1.3370

表 2　子波网络逼近结果

Table 2 Approximation result by wavelet network

函数	e^x	$\sin x$	$x+1$	x^2+1
迭代次数	1300	500	1300	1000
误差 err	0.0438	0.0573	0.0460	0.0285

表 3　BP 网络逼近结果

Table 3 Approximation result by BP network

函数	e^x	$\sin x$	$x+1$	x^2+1
误差	0.05	0.05	0.05	0.05
迭代次数(冲量算法)	4500	3500	5000	6500
迭代次数(一般算法)	13000	9000	14000	15000

4　结论

本文提出用于信号(函数)逼近的子波网络的初值设定法,大大加快了网络的收敛速度.通过对各种函数的实验以及对不同网络的对比,显示了此种方法的普遍性.由于此种子波网络对函数或信号逼近的精确性(可以收敛到很小的误差).我们可以展望它用于信号传输中,将大量采样数据的传输变为子波网络参数的传输,从而达到数据压缩的目的.

参考文献

1　焦成成.**神经网络的应用与实现**.西安:电子科技大学出版社,1992,237

2　Hurold H S. *Qptical Enginering*,1992,**31**(9):1907～1916

3　Zhang Q,Benveniste A. *IEEE transactions on Neural Networks*,1992,**3**(6):889～898

4　石卓尔,焦李成,保　铮.**子波神经网络**,1993 年学术大会论文集,1993,85～95

5　Rumelhart D E,Mccleland J L. *Exploration of Microstructure Cognition*, *Parallel Ditribution Processing*, MIT Press,1986

6　宋湘豫,李卫东,戚飞虎.*IEEE Int Conference on Neural Network and Signal Procedding*,1995,**1**:386～389

A NEW WAY OF SETTING INITIAL PARAMETER VALUE FOR FAST CONVERGENCE WAVELET NEURAL NETWORK USED FOR SIGNAL OR FUNCTION APPROXIMATION*

Song Xiangyu Chen Jianming Qi Feihu
(Optic-Fiber Technology Research Institute, Shanghai Jiao Tong University, Shanghai 200052, China)

Abstract A new way to set the initial values of the wavelet neural network's parameters was proposed in order to improve the convergence speed. Experiments on linear polynomial, exponent function, sin & cos function and a certain multi-stage simulation function show that the neural network has much faster convergence speed than wavelet networks which use other ways to set the initial parameter value and can be widely used for approximating many kinds of signals and functions. Discussion on this way's merits was given. Comparison was made between BP network and wavelet neural network. Experiment results are satisfactory.

Key words wavelet, wavelet transform, neural network, function approximation.

* The project supported by the National Climbing Foundation of China