

## 区间小波神经网络(I)——理论与实现\*

高协平<sup>1,3</sup> 张敏<sup>2,3</sup>

<sup>1</sup>(湘潭大学数学系 湘潭 411105)

<sup>2</sup>(清华大学计算机科学与技术系 北京 100084)

<sup>3</sup>(清华大学智能技术与系统国家重点实验室 北京 100084)

**摘要** 本文提出了前馈神经网络学习的一种新理论——区间小波神经网络,不同于以往工作的是本工作的主要特点有:(1)采用区间小波空间作为神经网络的学习基底空间,克服了以往神经网络基空间与被学习信号所属空间不匹配问题;(2)由于采用区间小波理论,克服了原来被学习信号为适应神经网络基空间而延拓所带来的不光滑性,使神经元数目得以节约,这在高维学习情形效果极为显著;(3)神经单元所用活性函数不再为同一个函数。

**关键词** 神经网络,小波,多尺度分析。

**中图法分类号** TP18

由于神经网络对复杂过程的高度非线性处理能力,它已得到蓬勃发展。其中,前馈神经网络因其结构简单及应用上的巨大成功而受到越来越多的重视,目前已成为从输入输出数据模拟非线性模型的普适学习手段。其应用几乎覆盖所有的工程实际领域,如自适应模式识别、实时语音重建、图象重构与数据压缩、实时分类、航空航天及机器人领域的大自由度系统控制等等。<sup>[1-3]</sup>另一方面,近年来迅速发展的小波多尺度理论作为一种新的逼近工具已呈现出强大的生命力,寻求小波理论在神经网络中的应用成为关注的热点。注意到小波分解与前馈神经网络学习的结构类似性,文献[4]~[6]提出了一种新的网络结构——小波神经网络,理论分析和模拟实例指出了对描述映射的令人惊讶的学习能力。对1,2维映射的学习表明,这种小波神经网络比单纯的小波分解可使用更少的基函数单元,更优于同样数目的 Sigmoid 单元活性函数的神经网络。

然而,这种小波神经网络有两个明显的不足:首先是对高维映射的学习所用神经元数目仍嫌过多,特别是在通常的实际情形中会不必要地浪费许多神经元;其次是随着维数  $d$  的增大,神经网络学习速度将减慢,即所谓“维数灾难”问题并没有解决。

众所周知,在绝大多数的实际问题中,被分析的信号(连续或离散)通常限制于一个区间或一个紧致有界空间上(如图象)。因此人们感兴趣的问题往往是囿于区间或紧致有界闭集上而非实轴或空间  $R^d$ 。我们假设被分析的信号  $f$  定义在区间上,不妨设为  $[0, 1]$ 。在文献[4]~[6]中,小波神经网络使用  $L^2(R)$  上的小波多尺度分析。从逼近论角度来看,使用  $L^2(R)$  上的小波神经网络来学习有限区间  $[0, 1]$  上的能量有限模拟信号  $f(x)$ ,存在逼近空间与信号空间  $L^2[0, 1]$  不匹配问题。为了使多尺度结构适合到有限区间  $[0, 1]$ ,传统的方案有如下3种:(1)零化扩张。通过置  $f(x) = 0, x \in [0, 1]$ ,将  $f(x) \in L^2[0, 1]$  扩张至整个实轴  $R$ ;(2)周期化。将  $f(x) \in L^2[0, 1]$  作周期1延拓至实轴  $R$ ;(3)镜像反射。将  $f(x) \in L^2[0, 1]$  作镜像延拓,即令  $g(x) = f(2n-x)$ ,当  $2n-1 \leq x \leq 2n$ ;  $g(x) = f(x-2n)$ ,当  $2n \leq x \leq 2n+1$ ;  $n \in Z$  为整数。

下面几点理由说明上述方案是有缺点的。首先,方案(1)(2)将使  $f(x)$  在端点  $x=0$  和  $x=1$  处人为地引入不连续性(除非  $f(x)$  在  $x \rightarrow 0$  和  $x \rightarrow 1$  时已经衰减到零或  $f(x)$  已经为周期的)。这种在端点引入的不连续性将导致端点附近小尺度下的衰减极慢的小波系数,尽管  $f(x)$  本身在  $[0, 1]$  上可能十分光滑。如果  $f(x)$  在  $[0, 1]$  上连续,则扩张方案(3)将使  $g(x)$  在  $R$  上连续,但通常会在整节点  $x=n \in Z$  上带来导数的不连续性。其次,上述扩张方案(1)~(3)将导致必

\* 本文研究得到国家自然科学基金、国家863高科技项目基金和国家攀登计划基金资助。作者高协平,1965年生,副教授,主要研究领域为计算智能,计算数学及应用软件。张敏,1935年生,教授,博士生导师,中国科学院院士,主要研究领域为人工智能,计算机应用技术。

本文通讯联系人,高协平,湘潭411105,湘潭大学数学系

本文1997-01-21收到原稿,1997-09-01收到修改稿

须使用“太多”的小波。设父小波  $\varphi(x)$  及母小波  $\psi(x)$  均有支撑集  $[0, 2N_0 - 1]$ ，在尺度  $j$  上，我们发现  $\langle f, \varphi_{jk} \rangle \neq 0$  (或  $\langle f, \psi_{jk} \rangle \neq 0$ ) (这里  $(\cdot, \cdot)$  为  $L^2$  内积,  $\varphi_{jk}, \psi_{jk}$  定义见后面) 将有  $2^j + 2N_0 - 1$  个小波; 而对  $[0, 1]$  上的逼近问题, 我们直观地感到只需使用  $2^j$  个小波, 即在 1 维情形的小波神经网络学习中, 我们实际上已浪费了  $2N_0 - 1$  个神经元, 在高维情形, 这种浪费更为可观。最后, 由于上述扩张通常引入了函数或导数的不连续性, 我们不能象往常那样期望借助小波系数描述函数的正则性。基于上述理由, 我们认为, 对实际情形中定义在有限区间上能量有限信号  $f$  的小波神经网络学习, 采用  $L^2(R)$  上的小波多尺度结构是不理想的。本文注意到  $L^2[0, 1]$  不再具有平移不变性, 通过重新定义  $L^2[0, 1]$  上的多尺度分析, 提出了另一种网络, 称为区间小波神经网络。其主要特色和结果是: (1) 通过重新定义  $L^2[0, 1]$  上的多尺度分析使小波多尺度结构与模拟信号空间  $L^2[0, 1]$  相匹配, 而非扩张模拟信号空间使之与  $L^2(R)$  小波多尺度结构匹配。正因为如此, 才有 (2) 克服了模拟信号延拓导致不连续性带来的神经元浪费问题, 特别在高维学习情形, 这种神经网络节约极为可观; (3) 神经单元采用不同的活性函数, 解决了  $L^2[0, 1]$  对平移的非不变量问题; 在另一篇论文“区间小波神经网络 I —— 性质与模拟”<sup>[15]</sup> 中我们将进一步有: (4) 证明了区间小波神经网络不存在高维学习的“维数灾难”, 即在  $d$  维情形神经网络学习速度不随  $d$  增大而减慢; (5) 模拟实例与理论相吻合, 证明了区间小波神经网络优于文献 [4]~[6] 提出的小波神经网络。

### 1 $L^2[0, 1]$ 上的多尺度分析 (MRA)

多尺度分析 MRA (multi-resolution analysis) 为构造小波基提供了一个统一的框架。<sup>[7-9]</sup> 本节简述了  $L^2(R)$  上的多尺度分析基本方法后, 重新定义了  $L^2[0, 1]$  上的 MRA, 并作多维拓广。

#### 1.1 $L^2(R)$ 上的 MRA

设  $\varphi(x) \in L^2(R)$  满足  $\int_{-\infty}^{+\infty} \varphi(t) dt = 1, \text{Supp} \varphi \subseteq \overline{\{x: \varphi(x) \neq 0\}} = [0, 2N_0 - 1], N_0$  为某正整数, 称  $\varphi(x)$  张成  $L^2(R)$  的一个正交多尺度分析 (OMRA)。如果存在增加闭子空间序列  $V_j \subset L^2(R)$ , 满足:

- (1)  $\dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots$ ;
- (2)  $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}, \bigcup_{j \in \mathbb{Z}} V_j = L^2(R)$ ;
- (3)  $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}, f(x) \in V_0 \Leftrightarrow f(x+1) \in V_0$ ;
- (4)  $\{\varphi(x-k)\}_{k \in \mathbb{Z}} \in V_0$  的一组正交基, 且有双尺度关系,  $\varphi(x) = \sum_{k=0}^{2N_0-1} h_k \sqrt{2} \varphi(2x-k)$ ;
- (5)  $\{h_k\}$  满足关系  $\sum_{k=0}^{2N_0-1} (-1)^k h_k \cdot k^a = 0, 0 \leq a \leq N_0 - 1$ ,

函数  $\varphi(x)$  称为该 OMRA 的父小波或尺度函数, 由离散镜像滤波器生成的函数  $\psi(x) = \sum_k (-1)^k \bar{h}_{1-k} \sqrt{2} \varphi(2x-k)$  称为母小波函数。记  $\varphi_{jk}(x) = 2^{j/2} \varphi(2^j x - k), \psi_{jk}(x) = 2^{j/2} \psi(2^j x - k), j, k \in \mathbb{Z}$ 。

定义  $W_j$  是  $V_j$  在  $V_{j+1}$  中的正交补空间, 即  $V_{j+1} = V_j \oplus W_j$ , 则可知  $\{\varphi_k, j, k \in \mathbb{Z}\}, \{\psi_k, j, k \in \mathbb{Z}\}$  分别构成  $V_j, W_j$  的正交基底, 并且我们有如下重要的结论:

- (1)  $L^2(R) = \bigcup_{j \in \mathbb{Z}} V_j = \bigoplus_{j \in \mathbb{Z}} W_j = V_{j_0} \oplus W_j, \forall j_0 \in \mathbb{Z}$ ;
- (2) 对  $\forall f(x) \in L^2(R), \forall \epsilon > 0, \exists M > 0$ , 使  $\|f(\cdot) - \sum_{k \in \mathbb{Z}} \varphi_{jk}(\cdot) \int_{-\infty}^{+\infty} f(t) \varphi_{jk}(t) dt\|_{L^2(R)} < \epsilon$ 。

#### 1.2 $L^2[0, 1]$ 上的 MRA

我们知道,  $L^2(R)$  情形下小波基的构造基于 Fourier 技巧, 即依赖于平移不变性。如果将函数限制于一有限区间, 如  $[0, 1]$ , 那么在区间端点处建立了一种非齐性, 它破坏了平移不变性。因此  $L^2[0, 1]$  不再保持平移不变性。为了构造  $L^2[0, 1]$  的小波基组, 多尺度分析必须重新定义。特别地, 我们仅仅考虑尺度  $j \geq j_0 \geq 0$ , 使尺度函数和小波函数的支撑限于区间  $[0, 1]$  内, 即考虑一个逼近子空间半序列  $V_{j_0} \subset V_{j_0+1} \subset \dots \rightarrow L^2[0, 1]$ , 对某些  $j_0 \geq 0$ 。

由 1.1 节,  $L^2(R)$  上 MRA 定义中条件 (5), 我们知道所有次数  $\leq N_0 - 1$  的多项式  $P_{N_0-1}$  可由函数  $\{\varphi_k; k \in \mathbb{Z}\}$  的某线性组合获得。由于这个事实与小波的许多逼近性质直接相关, 任何区间小波的构造均应保持此性质。对  $L^2[0, 1]$  上的构造, 我们将以此为出发点。首先, 我们构造空间  $V_j[0, 1]$ , 它产生  $L^2[0, 1]$  上的正交多尺度分析。

设  $j_0$  是满足如下条件的最小正整数:  $j_0 = \min\{j: 2^{j-1} \geq N_0\}$ , 定义序列  $\{m_i\}_{i \geq 0}, \{M_i\}_{i \geq 0}$  为:

$$m_0 = \sqrt{2}, m_i = \sum_{k=0}^{2N_0-1} h_k \cdot k^i; \quad M_0 = 1, M_i = \frac{1}{(2^i - 1)\sqrt{2}} \sum_{j=0}^{i-1} \binom{i}{j} m_{i-j} M_j, i \geq 1$$

对  $0 \leq a \leq N_0 - 1, j \geq j_0$ , 我们定义:

$$X_{j,a}^g(x) = \sum_{k=-2N_0+2}^0 \left( \sum_{i=0}^a \binom{a}{i} k^i M_{a-i} \right) \varphi_k(x) |_{[0,1]}, \quad Y_{j,a}^g(x) = \sum_{k=-2N_0+1}^{-1} \left( \sum_{i=0}^a \binom{a}{i} k^i M_{a-i} \right) \varphi_k(x-1) |_{[0,1]}$$

定义伸缩算子  $T_n g(x) = 2^{n/2} g(2^n x)$ , 利用数学归纳法容易证明 Gram-Schmidt 正交化过程保持  $T_n$  意义下的伸缩性. 这样, 利用 Gram-Schmidt 过程, 分别正交化  $X_{j,a}^g(x)$  和  $Y_{j,a}^g(x)$  获得  $\varphi_k^g(x)$  和  $\varphi_k^g(x)$ . 记  $\varphi_k^g(x) = 2^{j/2} \varphi_k^g(2^j x), \varphi_k^g(x) = 2^{j/2} \varphi_k^g(2^j x)$ , 定义函数组  $\{\varphi_k^g(x)\}$ :

$$\varphi_k^g(x) = \begin{cases} \varphi_k^g(x), & k = -N_0 + 1(1)0 \\ \varphi_k(x), & k = 1(1)2^j - 2N_0 \\ \varphi_k^{2^j - 2N_0 + 1}(1)2^j - N_0, & k = 2^j - 2N_0 + 1(1)2^j - N_0 \end{cases}$$

那么我们有下述定理.

**定理 1.** 设  $j \geq j_0 \geq \log_2 N_0 + 1$ , 则函数组  $\{\varphi_k(x)\}_{k=-N_0+1}^{2^j - N_0}$  是空间  $V_j[0,1]$  的一组正交基, 且  $V_j[0,1]$  构成一个增加族序列.  $V_{j_0}[0,1] \subset V_{j_0+1}[0,1] \subset \dots \rightarrow L^2[0,1]$ . 该定理的证明详见文献[10].

由此定理, 我们容易看到,  $\dim V_j[0,1] = 2^j$ . 且对任给  $f(x) \in L^2[0,1]$ , 对  $\forall \epsilon > 0$ , 存在  $M \geq j_0$  足够大, 使

$$\|f(\cdot) - \sum_{k=-N_0+1}^{2^M - N_0} \int_0^1 f(x) \varphi_k(x) dx \cdot \varphi_k(\cdot)\|_{L^2[0,1]} < \epsilon \quad (1)$$

其次, 如果记  $W_j[0,1], j \geq j_0$  为  $V_j[0,1]$  在  $V_{j+1}[0,1]$  中的正交补空间, 完全类似地, 可选取  $2^j$  个函数  $\{\psi_k(x)\}_{k=0}^{2^j-1}$  构成  $W[0,1]$  的正交基. 从本文的实际需要出发, 这里省略其求解过程与表示, 有兴趣者详见文献[10].

### 1.3 多维推广

对  $d$  维情形, 考虑超立方体  $[0,1]^d$ , 我们可由 1 维情形的尺度函数及小波函数利用张量积产生  $d$  维正交多尺度分析. 为简单计, 采用如下记号与约定:  $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$  定义为  $L^2[0,1]$  上的内积;  $\langle f, g \rangle = \int_0^1 \dots \int_0^1 f(x_1, \dots, x_d)g(x_1, \dots, x_d)dx_1 \dots dx_d$  定义为  $L^2([0,1]^d)$  上的内积;  $\|f\|_{L^2} = \langle f, f \rangle^{1/2}$  为  $L^2([0,1]^d)$  上的  $L^2$  范数.  $J_n = \{k, -N_0 + 1$

$\leq k \leq 2^n - N_0\}; L_n = \{k, 0 \leq k \leq 2^n - 1\}; U = \{u_1, \dots, u_d\}, D = \{1, \dots, d\}, \prod_{i=1}^d 1 = 1$ . 略去超立方体  $[0,1]^d$  记号, 简以  $V_n,$

$W_n$  分别表示  $L^2([0,1]^d)$  上尺度  $n$  下的尺度空间与小波空间, 则直接地, 我们知道下述结论成立:  $\{\prod_{i=1}^d \varphi_{k_i}(x_i); k_i \in J_n\}$

构成  $V_n$  的一组正交基;  $\{\prod_{i=1}^d \varphi_{k_i}(x_{q_i}) \cdot \prod_{l=i+1}^d \psi_{q_l}(x_{u_l}); 0 \leq i \leq d-1, u \in D, k_i \in J_n; q_i \in L_n\}$  构成  $W_n$  的一组正交基;

且  $L^2([0,1]^d) = V_n \oplus W_n = \bigcup_{j \geq j_0} V_j, n \geq j_0$ . 因此, 对任意  $f(x_1, x_2, \dots, x_d) \in L^2([0,1]^d), \forall \epsilon > 0$ , 存在  $M \geq j_0$  足够大, 使

$$\|f(x_1, x_2, \dots, x_d) - \sum_{\substack{k_i \in J_M \\ i=1(1)d}} \langle f, \prod_{j=1}^d \varphi_{k_j} \rangle \prod_{i=1}^d \varphi_{k_i}(x_i)\|_{L^2} < \epsilon.$$

## 2 区间小波神经网络

### 2.1 结构

为简单计, 我们首先考虑 1 维信号(单输入)情形. 设  $f(t) \in L^2[0,1]$  是任意模拟信号. 给定映射  $f$  下的输入输出对样本集合:  $T_n = \{(t_i, f(t_i)); i=1(1)n; 0 \leq t_i \leq 1\}$ . 我们寻求对未知映射  $f$  的好的逼近. 从  $L^2[0,1]$  上的多尺度分析, 我们知道,  $f(t)$  可以通过选择足够大的  $M$  而在  $V_M$  中被任意好地逼近:

$$f(t) \approx \sum_{i \in J_M} \langle f, \varphi_{k_i} \rangle \varphi_{k_i}(t) \quad (2)$$

在通常的神经网络中, 每个单元结点有如下的输入输出关系:

$$S_{\text{out}}^j = g_j \left( \sum_m \omega_m S_{\text{in},m}^j + I_j \right)$$

其中  $S_{\text{in},m}^j$  是结点  $j$  的第  $m$  个输入,  $\omega_m$  是权重,  $I_j$  是阈值,  $g_j(\cdot)$  是结点  $j$  的非线性活性函数. 通常, 对所有结点  $j$ ,

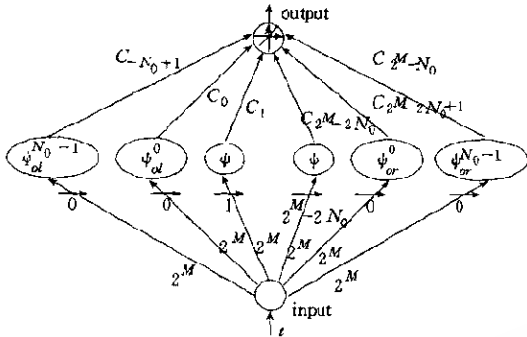


图1 区间小波神经网络的实现

$g_j(\cdot) \equiv g(\cdot)$ . 如在 MLP 模型中常用的 Sigmoid 函数. 因此, 式(2)的逼近可由一个三层前馈神经网络实现, 如图 1 所示. 对一个给定的  $M$ , 图 1 描述的区间小波神经网络实现一个函数  $g(t) = \sum_{k \in J_M} C_k \varphi_{Mk}(t)$ . 通过学习, 合适地选取连接权  $C_k$

可很好地实现对  $f(t)$  的逼近. 高维情形的推广是直接的.  $\{\prod_{j=1}^d \varphi_{Mk_j}(x_j); k_j \in J_M\}$  是  $L^2([0, 1]^d)$  的一组分离尺度函数, 它产生一个正交多尺度结构, 对感兴趣的模拟信号  $f(x_1, \dots, x_d) \in L^2([0, 1]^d)$ , 可在合适的尺度  $M$  下被  $V_M$  任意好地逼近, 即  $d$  维区间小波神经网络可由映射  $g(x_1, \dots, x_d)$  实现, 这里

$$g(x_1, \dots, x_d) = \sum_{k_1, \dots, k_d \in J_M} C_{k_1, \dots, k_d} \prod_{j=1}^d \varphi_{Mk_j}(x_j) \quad (3)$$

## 2.2 训练过程

### 2.2.1 训练算法

定义均方根训练误差 ( $x = (x_1, \dots, x_d)$  为坐标变量)  $e_n(f, g) = \frac{1}{n} \sum_{i=1}^n (f(x^i) - g(x^i))^2$ . 如果训练样本集  $T_n$  是有效的, 则对信号  $f(x)$  的逼近解  $g(x)$ , 即  $(\tilde{C}_{k_1, \dots, k_d})_{k_j \in J_M}$  可由极小化均方根误差  $e_n(f, g)$  获得.

$$(\tilde{C}_{k_1, \dots, k_d})_{k_j \in J_M} = \underset{(C_{k_1, \dots, k_d})_{k_j \in J_M}}{\operatorname{argmin}} e_n(f, g) \quad (4)$$

方程(4)可用通常的最优化方法求解, 如迭代梯度下降过程.

$$\tilde{C}_{k_1, \dots, k_d}^{(p)} = \tilde{C}_{k_1, \dots, k_d}^{(p-1)} - \lambda \frac{\partial e_n(f, g)}{\partial C_{k_1, \dots, k_d}} \Big|_{C_{k_1, \dots, k_d} = \tilde{C}_{k_1, \dots, k_d}^{(p-1)}}, \quad p \geq 1 \quad (5)$$

这里  $\lambda$  是迭代步长.

式(4)也可通过直接对  $e_n(f, g)$  求偏导数用极值法求解:  $\frac{\partial e_n(f, g)}{\partial C_{k_1, \dots, k_d}} = 0$ , 它最终导致一个线性代数方程组. 由于尺度函数是正交的, 对足够大的  $n$ , 该线性代数方程组有唯一解, 但系数矩阵一般而言是非稀疏的, 从而求解代价是昂贵的. 从实际应用角度出发, 我们宁愿推荐采用迭代梯度下降过程式(5)求解, 并使用迭代初始权值猜测.

$$\tilde{C}_{k_1, \dots, k_d}^{(0)} = \frac{1}{n} \sum_{i=1}^n f(x^i) \prod_{j=1}^d \varphi_{Mk_j}(x_j^i) \quad (6)$$

### 2.2.2 实用注记

(1) 关于网络规模设计, 即隐层单元结点数. 因为尺度空间  $V_M$  的维数精确地为  $2^M$ , 因此对给定的  $M$ , 隐层单元结点数不超过  $2^{Md}$  个. 考虑到网络的泛化能力, 隐层单元数应控制在绝对不超过训练样本数范围内, 即  $2^{Md} \leq n$ . 这样, 我们可以给出尺度  $M$  的一个界估计:  $j_0 \leq M \leq \frac{1}{d} \log_2 n$ . 通常, 上述  $M$  的界限长度较小, 我们在此界限中选取一个  $M$  (如取界限平均), 利用前述算法训练网络, 并求出均方根误差, 在允许精度内停止, 否则  $M$  加 1 进行训练.

(2) 对高维情形, 若训练样本数据非一致分布于  $[0, 1]^d$  或训练数据稀疏以至不能“张满”(Fill Out)函数域  $[0, 1]^d$ , 如文献[6]所建议一样, 我们可以借助于主分量分析回归 PCAR (principle-component analysis regression) 或投影追赶回归 PPR (projection pursuit regression) 等技巧, 利用一些神经网络的合成逼近信号  $f$ .

### 参考文献

- 1 Poggio T, Girosi F. Networks for approximation and learning. In: Proceedings of the IEEE, 1990, 78(9): 1481~1497
- 2 Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Networks, 1989, 2(5): 359~366
- 3 Hwang J N, Lay S R, Maechler M et al. Regression modeling in back-propagation and projection pursuit learning. IEEE Transactions on the Neural Networks, 1994, 5(3): 342~353
- 4 Pati Y C, Krishnaprasad P S. Analysis and synthesis of feedforward neural networks using discrete affine wavelet transformations. IEEE Transactions on the Neural Networks, 1993, 4(1): 73~85
- 5 Zhang Qing-hua, Benveniste A. Wavelet networks. IEEE Transactions on the Neural Networks, 1992, 5(6): 889~898

- 6 Zhang Jun, Walter G G, Miao Y *et al.* Wavelet neural networks for function learning. *IEEE Transactions on the Signal Processing*, 1995, 43(6):1485~1497
- 7 Mallat S G. Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on the Acoust. Speech, Signal processing*, 1989, 37(12):2091~2110
- 8 Chui C K. An introduction to wavelets. New York: Academic Press, 1992.
- 9 Daubechies I. Orthonormal bases of compactly supported wavelets. *Communications in Pure and Applied Mathematics*, 1988, 41(8):909~996
- 10 Andersson L, Hall N, Jawerth B *et al.* Wavelets on closes subsets of the real line. In: Schumaker L L, Webb G eds. *Recent Advance in Wavelet Analysis*. Boston: Academic Press, 1994. 1~61

## Interval-wavelets Neural Networks(I) —— Theory and Implements

GAO Xie-ping<sup>1,3</sup> ZHANG Bo<sup>2,3</sup>

<sup>1</sup>(Department of Mathematics Xiangtan University Xiangtan 411105)

<sup>2</sup>(Department of Computer Science and Technology Tsinghua University Beijing 100084)

<sup>3</sup>(National Key Laboratory of Intelligent Technology and Systems Tsinghua University Beijing 100084)

**Abstract** In this paper, an interval wavelets neural networks is proposed as an alternative to feedforward neural network for approximating arbitrary nonlinear functions. Different from the past ones, the present work has main characteristics as follows: (1) using interval wavelet space as the basic learning space of neural networks, the authors have solved the problem in which basic space does not match the space of learnt signals. (2) As interval wavelets theory is used, the authors have overcome the discontinuity problem caused by enlarging learnt signals in order to adapt basic space of neural network. The number of neurons is decreased, which is greatly notable in high-dimensional situations. (3) The activity functions of neurons are not the same functions.

**Key words** Neural network, wavelets, multiresolution analysis.

**Class number** TP18