



# 一种用于函数学习的小波神经网络

吕柏权 李天铎 吕崇德 刘兆辉

(清华大学热能工程系 北京 100084)

**摘要** 在非线形系统辨识中,系统输入往往是多变量的.小波处理此类问题则比较复杂.结合神经网络形式和小波特点建立一种新型的网络,可简单有效地解决网络多输入问题.同时给出此网络可以逼近任意连续函数的数学证明.并通过实例验证了此方法的正确性.

**关键词** 参数辨识,小波网络,BP神经网络.

## 1 引言

小波分析是近年来迅速发展起来的新学科.它广泛地用于信号分析、故障诊断等方面.其最大特点是它在时域-频域同时具有良好的局部化性质.另外小波还能把  $L^2(\mathbb{R})$  空间分成闭子空间  $W_J, J \in \mathbb{Z}$  的直和形式(尺度分析).目前利用此特点进行函数逼近也倍受关注<sup>[1-3]</sup>.但对多维情况,小波处理此问题比较复杂.在实时单输入-单输出的控制系统参数辨识中,往往又要求网络输入为多变量的.为此本文利用神经网络形式,以小波函数做激励函数从而简单而有效地解决网络多输入问题,同时给出此网络可以逼近任意连续函数的数学证明,并通过实例验证了此方法的正确性.

## 2 小波变换及学习算法

首先考虑一维的情况,记  $\langle f, g \rangle = \int_{\mathbb{R}} f(x)g(x)dx$   $\|f\| = \langle f, f \rangle^{1/2}$ ,  $\Psi^{(a,b)}(x) = \frac{1}{\sqrt{a}}\Psi(\frac{x-b}{a})$ ,  $T(a,b) = \langle f, \Psi^{a,b} \rangle$ .

其中  $\Psi$  为母小波,即

$$C_{\Psi} = \int_{-\infty}^{\infty} |\hat{\Psi}(\omega)|^2 \omega^{-1} d\omega < \infty. \tag{1}$$

**定义 1.** 函数  $f(x) (\in L^2(\mathbb{R}))$  空间的连续小波变换为  $T(a,b)$ . 其反变换为

$$f(x) = \frac{1}{C_{\Psi}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T(a,b) \Psi^{(a,b)}(x) \frac{dadb}{a^2}.$$

当  $a=2^j, b=kb_02^j (j, k \in \mathbb{Z})$  时,反变换离散化为

$$f(x) = \sum_{j,k \in \mathbb{Z}} \langle f, \Psi_{j,k} \rangle \Psi^{j,k}(x) = \sum_{j,k \in \mathbb{Z}} \langle f, \Psi^{j,k} \rangle \Psi_{j,k}(x). \tag{2}$$

其中  $\Psi_{j,k}$  是  $\Psi^{j,k}$  的共轭基,  $\Psi^{j,k}(x) = 2^{-j/2} \Psi(2^{-j}x - k b_0)$ .

为了能再现  $f(x)$  还必须满足如下条件, 即框架(Frame)

$$A \|f\|^2 \leq \sum_{j,k \in \mathbb{Z}} |\langle f, \Psi^{j,k} \rangle|^2 \leq B \|f\|^2 \quad \text{且 } 0 \leq A \leq B < \infty,$$

(2)式可近似为  $\hat{f}(x) = \sum_{K=1} C_K \Psi(a_K x - b_K)$ , 这可以用三层 NN 网络实现. 当维数为  $d$  时判断  $\Psi_d(x)$  满足框架条件比较困难<sup>[1,2]</sup>, 同样,  $\Psi_d(x)$  为小波函数时, 可以用  $\Psi_d(x) = \Psi_d(x_1, x_2, \dots, x_d) = \prod_{j=1}^d \Psi_d(x_j)$  表示. 此表示法在  $x$  的维数比较高时计算量比较大. 本文用  $\hat{f}_1(x) = \sum_{K=1} C_K \Psi(\alpha_K^T x - b_K)$  来近似  $f(x)$ , 其中  $\alpha_K^T$  为行向量,  $X$  为列向量, 现在讨论  $\Psi(x)$  是怎样的函数时用此近似  $f(x)$  的误差在允许范围内.

**定义 2**<sup>[4]</sup>. 设函数  $g(x) \in L^p_{loc}(R^1)$ , 若线性组合  $\sum_{i=1}^n c_i g(\lambda_i x + \theta_i)$  在任意  $L^p[a, b]$  中稠密, 则称  $g$  是一个  $L^p$ -Tauber-Wiener 函数, 简记  $g \in (L^p TW)$  其中  $C_i, \theta_i$ , 及  $\lambda_i \neq 0$  均为实数,  $i=1, \dots, N$ . 若  $g$  为一个一元函数(连续的或不连续的)且  $\sum_{i=1}^n c_i g(\lambda_i x + \theta_i)$  在任意  $C[a, b]$  中稠密, 则称  $g$  是一个 C-Tauber-Wiener 函数. 简记  $g \in (CTW)$ .

**引理 1.** 设  $K$  为  $R^n$  中任一个紧集,  $V$  为  $L^p(K)$  一个紧集,  $g \in (L^p TW (1 \leq p \leq \infty))$ , 则对任意  $\epsilon > 0$  存在不依赖  $f$  的一正整数  $N$ , 实常数  $\theta_i$  以及向量  $y_i \in R^1$ , 和依赖于  $f$  的常数  $c_i(f), i=1, \dots, N$ , 使  $\|f(x) - \sum_{i=1}^N C_i g(y_i^T x + \theta_i)\|_{L^p(K)} < \epsilon$  对一切  $f \in V$  成立, 且所有  $c_i(f)$  都是定义在  $V$  上的连续泛函.

**定理 1.** 设  $K$  为  $R^n$  中任一个紧集,  $V$  为  $L^2(K)$  中的一个紧集, 对于框架函数  $\Phi \in L^2(IR)$ , 则对于任意  $\epsilon > 0$  存在一个正整数  $N$ , 实常数  $b_i$ , 向量  $a_i \in R^1$ , 均不依赖于  $f \in V$ , 以及依赖于  $f \in V$  的常数  $c_i(f)$ , 使  $\left| f(x) - \sum_{i=1}^N c_i \phi(a_i^T x - b_i) \right|_{L^2(K)} < \epsilon$  对一切  $f \in V$  成立. 所有  $c_i(f), i=1, \dots, N$  都是定义在  $V$  上的线性连续泛函.  $\hat{\phi}$  为  $\phi$  的付里叶变换.

证. 由引理 1 知道, 只要  $\hat{f}(x)$  (在任意  $L^p[a, b]$  中稠密(其中  $x \in R^1$ ), 则对任意  $f(x) \in V$  ( $V$  为  $L^p(K)$  一个紧集,  $K$  为  $R^n$  中任一个紧集), 都可以用  $\hat{f}_1(x)$  逼近(其中向量  $x \in R^1$ ), 由于  $\Psi$  是框架, 则保证了  $\hat{f}(x)$  在任意  $L^2(IR)$  中稠密(其中  $x \in R^1$ ), 也就是说用  $\hat{f}_1(x)$  是可以以任意精度逼近  $f(x)$ . 证毕.

**定理 2.** 设  $K$  为  $r^n$  中任一紧集,  $V$  为  $C(K)$  中  $R$  一个紧集, 对于框架函数  $\Phi \in L^2(IR)$ , 则对于任意  $\epsilon > 0$  存在一个正整数  $N$ , 实常数  $b_i$ , 向量  $a_i \in R^1$ , 它们均不依赖于  $f \in V$ , 以及依赖于  $f \in V$  的常数  $c_i(f)$  使  $\left| f(x) - \sum_{i=1}^N C_i \phi(a_i^T x - b_i) \right|_{C(K)} < \epsilon$  对一切  $f \in V$  成立. 所有  $c_i(f), i=1, \dots, N$  都是定义在  $V$  上的线性连续泛函.  $\hat{\phi}$  为  $\phi$  的付里叶变换.

**3 仿真结果**

本文以  $y'' = \sin(y') - (1 + 1.5y') * y^2 + (1 + 0.2y)u$  为例进行仿真. 仿真步长为  $h =$



0.01s). 每 5 倍  $h$  作为一个学习点, 共 100 点, 每 100 点作为一个学习周期重新开始学习.  $u=1.0$ , 其权初值为  $(-0.01, 0.01)$  的均匀分布;  $x = \{y(n-1), y(n-2), u(n)\}$  为网络的输入;  $y$  为网络的输出; 中间结点选为 200. 激励函数分别为

$$\Psi_1(x) = \frac{8}{1 + e^{-x}} - 4 \text{ (BP 网络)},$$

$$\Psi_2(x) = \frac{1}{\sqrt{2\pi}}(1 - x^2)e^{-x^2/2} \text{ (小波网络)}.$$

易验证小波  $\Psi_2(x)$  满足框架条件. 终止条件为

$$E_1 = \sum_{i=1}^{100} |y(i) - \hat{y}(i)|^2.$$

当激励函数为  $\Psi_1(x)$  且学习 7500 个周期时  $E_1$  才小于 0.064 (学习速率  $\eta=0.002$ ). 当激励函数为  $\Psi_2(x)$  且学习 440 个周期时,  $E_1$  已小于 0.008 (学习速率  $\eta=0.07$ ). 可见效果明显. 其仿真结果如图 1 和图 2 所示 (横轴表示采样点, 换算成时间秒  $0.05 \times$  采样数).

本文给出了一种小波网络逼近任意函数方法的证明, 从而为用这种小波网络逼近函数时提供理论根据. 并通过仿真验证此方法的正确性.

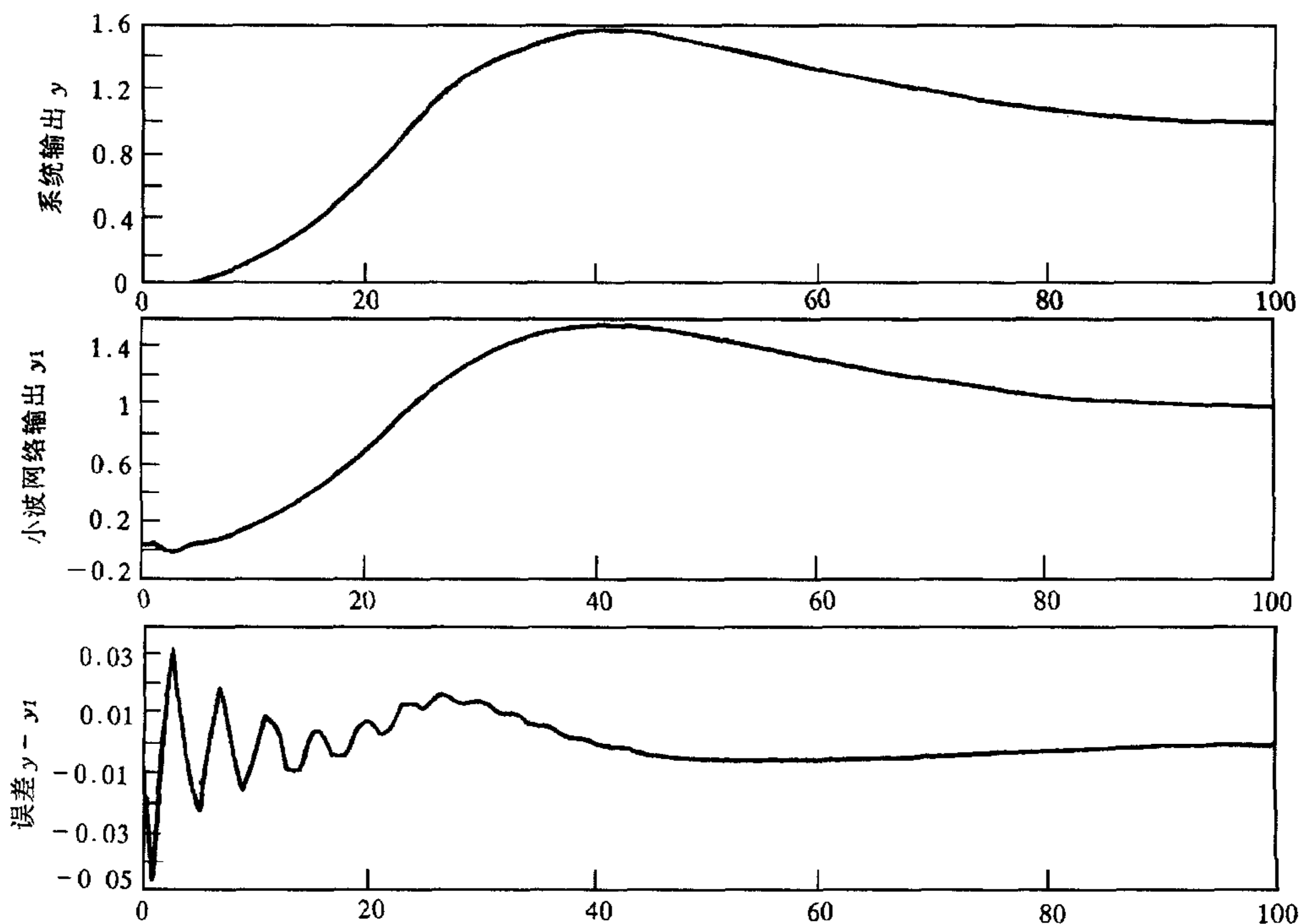


图 1 小波网络逼近结果

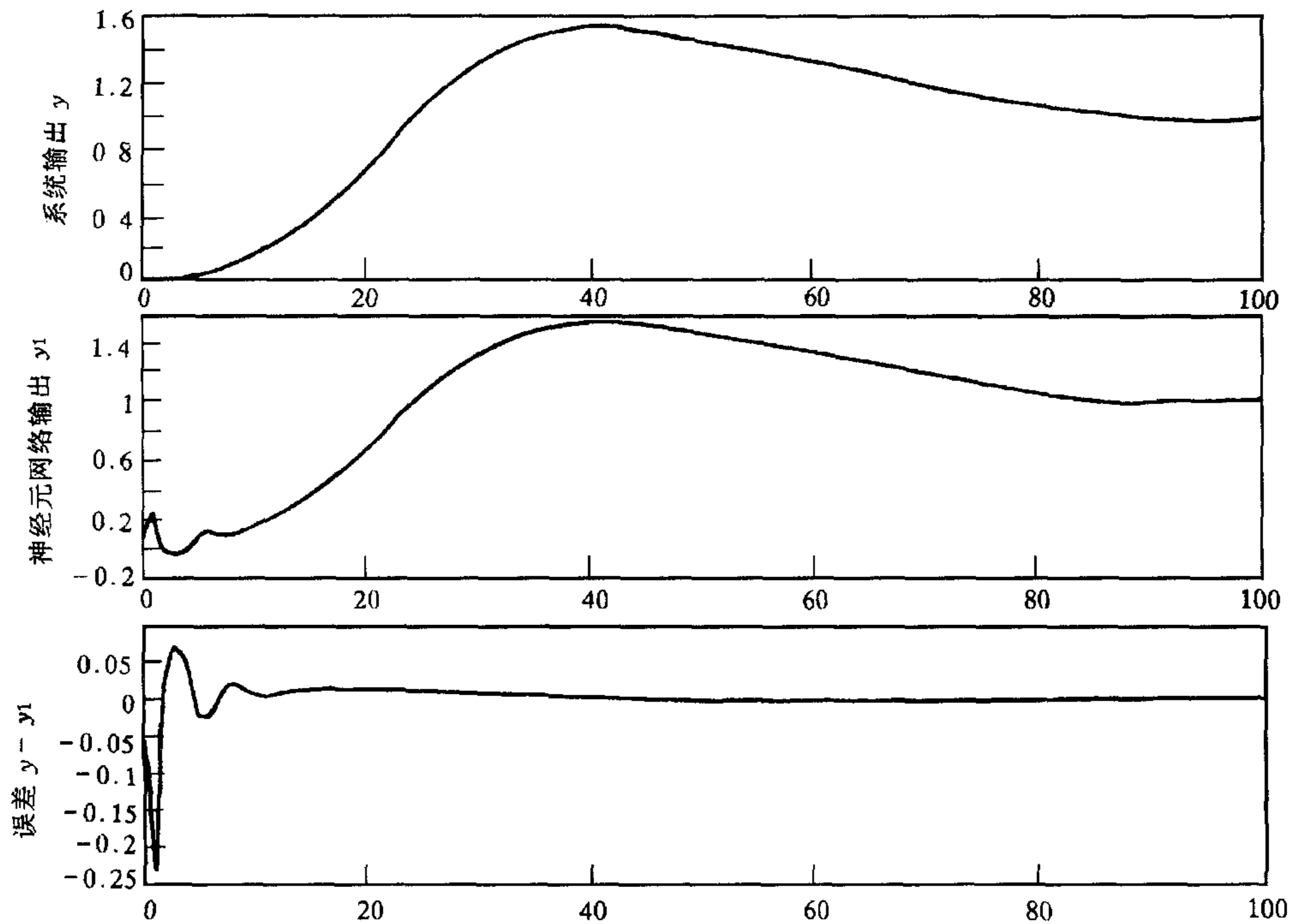


图 2 神经网络逼近结果

## 参 考 文 献

- 1 Delyon B, Juditsky A, Benvenise A. Accuracy analysis for wavelet Approximations. *IEEE Transactions on Neural Networks*, 1995, 6(2): 332-348
- 2 JunZhang, Gilbert G Walter, Yubo Miao et al. Wavelet Neural Networks for function learning. *IEEE Transactions on Signal Processing*, 1995, 43(6): 1485-1496
- 3 倪先锋, 陈宗基, 周绥平. 基于神经网络的非线性学习控制研究. *自动化学报*, 1993, 19(3): 307-314
- 4 陈天平. 神经网络及其在系统识别应用中的逼近问题. *中国科学学报(A 辑)*, 1994, 24(1): 1-7
- 5 秦前清, 杨宗凯. 实用小波分析. 西安: 西安电子科技大学出版社, 1994

## WAVELET NEURAL NETWORK FOR FUNCTION LEARNING

LÜ BAIQUAN LI TIANDUO LÜ CHONGDE LIU ZHAOHUI

(Department of Thermal Engineering, Tsinghua University, Beijing 100084)

**Abstract** In identification of nonlinear system, the network inputs usually are multivariables. The approach using wavelets alone to resolve the problem is complex. Therefore, the paper establishes a new network method by combining neural network and wavelet characteristics to handle the multivariable problem simply and effectively. This paper also gives the mathematical verification that the network can approximate any kind of functions. The simulation results of an example indicate the validity of the method.

**Key words** Parameters identification, wavelet network, BP neural network.