



用含动态隐层的前馈网辨识 非线性系统¹⁾

鲍晓红 贾英民

(北京航空航天大学第七研究室 北京 100083)

摘要 用含动态隐层的前馈网对非线性系统建立全局成立的输入输出模型,证明了这种辨识结构的可行性,网络学习算法为动态 BP 法.

关键词 前馈网,辨识,动态 BP 算法.

1 引 言

人们在用神经网络辨识非线性系统方面已做了大量工作,这些工作大都建立在非线性动态系统的输入输出模型上^[1],并用增加了输出时延反馈的多层前馈网(RFNN)来逼近它,效果显著.但值得注意的是这种输入输出模型是从离散非线性状态方程推出的局部成立的模型^[1],对原状态方程中的输入输出而言在全局上并不一定有此关系,针对这个问题,本文提出用含动态隐层的前馈网(RHNN)直接用状态方程描述的非线性系统的输入输出关系建立网络模型,学习算法为动态 BP 法,并在仿真中对此种系统的输入输出关系分别用 RFNN 和 RHNN 进行了辨识比较,结果表明后者能更准确地辨识这种系统.

2 用神经网络对非线性系统建模

一类广泛应用的非线性系统可以在大范围内(全局)用如下模型表示:

$$\mathbf{x}(k+1) = f(\mathbf{x}(k), \mathbf{u}(k)), \mathbf{y}(k) = h(\mathbf{x}(k), \mathbf{u}(k)). \quad (1), (2)$$

其中 $\mathbf{x}(k) \in X \subset R^n$ 为系统内部状态; $\mathbf{u}(k) \in U \subset R^r$ 为系统输入; $\mathbf{y}(k) \in Y \subset R^m$ 为输出; f, h 为连续函数.

设系统(1),(2)满足如下假设:

H1) $\forall \mathbf{u} \in U, \mathbf{x}_0 \in X$ 对于有限的 M (正整数),恒有 $\|\mathbf{x}(M)\| + \|\mathbf{y}(M)\| < \infty$,即系统是稳定的;

H2) f, h 满足 Lipschitz 条件,即系统的解是唯一的.

1) 国家自然科学基金资助项目.

定理1. 对满足 H1), H2) 的非线性系统(1)和(2), 可以用一个带动态隐层的四层前

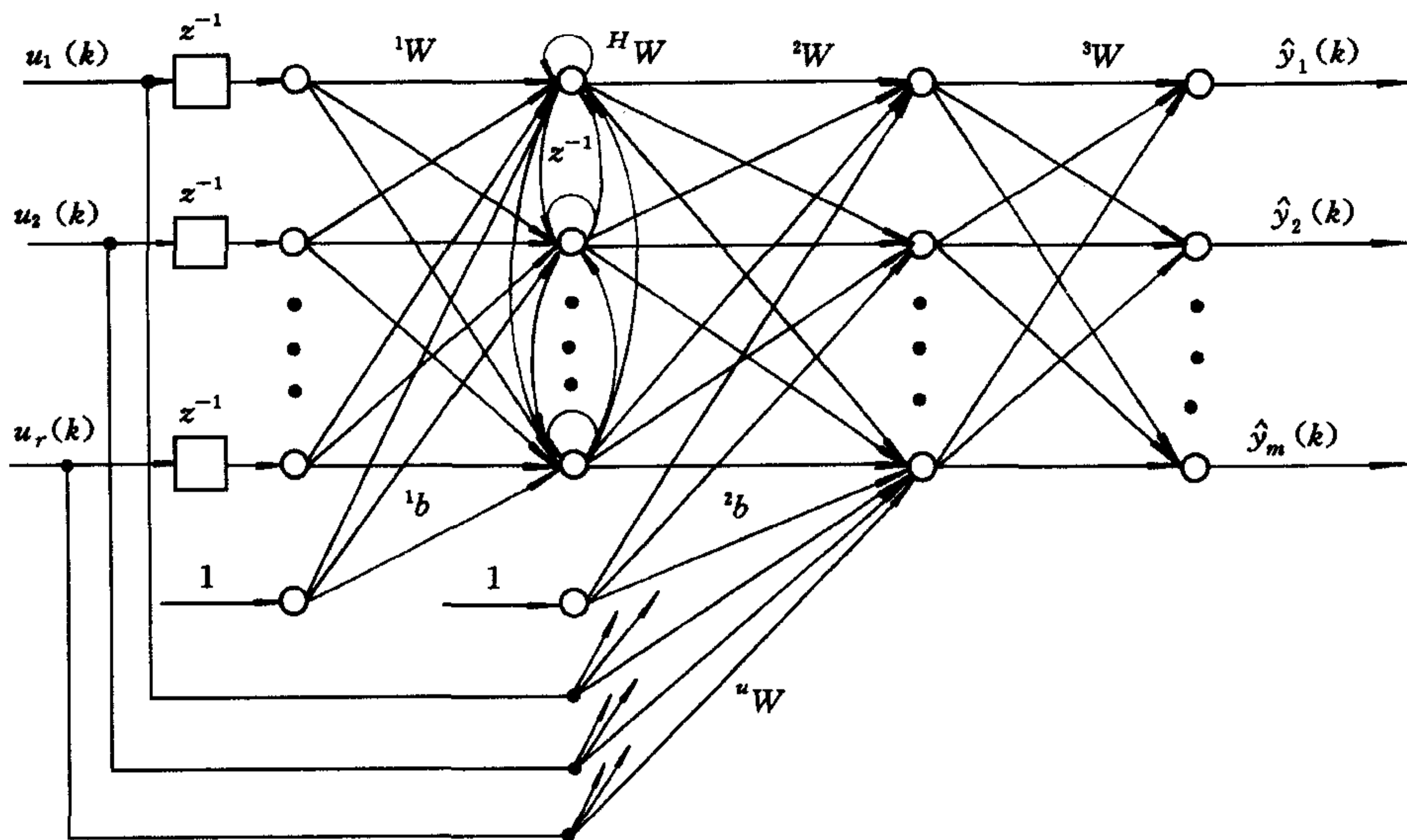


图1 RHNN 示意图

馈网(如图1所示)在 $k \in [0, M]$ 内以任意精度逼近其输入输出特性, 即对 $\forall \epsilon > 0$, 使

$$\max_{k \in [0, M]} \|y(k) - \hat{y}(k)\| < \epsilon. \tag{3}$$

证明. 用两个三层前馈网 NN_f, NN_h 分别逼近 $f(\cdot), h(\cdot)$. 对 NN_f , 其参数 $W_f = \{^1W_f, ^2W_f, ^1b\}$, 其中 $^1W_f \in R^{n_1 \times (n+r)}, ^2W_f \in R^{n \times n_1}, ^1b \in R^{n_1}$. 对 NN_h , 其参数 $W_h = \{^1W_h, ^2W_h, ^2b\}$, $^1W_h \in R^{n_2 \times (n+r)}, ^2W_h \in R^{m \times n_2}, ^2b \in R^{n_2}$, 则有

$$\begin{aligned} \hat{x}(k+1) &= ^2W_f \cdot ^1O(k+1), \\ ^1O(k+1) &= g \left[\begin{matrix} ^1W_{f\hat{x}} & ^1W_{fu} & ^1b \end{matrix} \cdot \begin{matrix} \hat{x}(k+1)z^{-1} \\ u(k) \\ 1 \end{matrix} \right]. \end{aligned} \tag{4}$$

其中 $g(a) = \frac{1}{1 + \exp^{-a}}$, 若 $\mathbf{x} = (x_1, \dots, x_n)^T, g(\mathbf{x}) = (g(x_1), \dots, g(x_n))^T$,

$$\begin{aligned} ^1O(k+1) &= g(^1W_{f\hat{x}} \cdot \hat{x}(k+1)z^{-1} + ^1W_{fu} \cdot u(k) + ^1b) \\ &= g(^1W_{f\hat{x}} \cdot ^2W_f \cdot ^1O(k) + ^1W_{fu} \cdot u(k) + ^1b). \end{aligned}$$

令 $^1W_{f\hat{x}} \cdot ^2W_f = ^HW$, 则有

$$^1O(k+1) = g(^HW \cdot ^1O(k) + ^1W_{fu} \cdot u(k) + ^1b). \tag{5}$$

又

$$\begin{aligned} \hat{y}(k+1) &= ^2W_h g \left[\begin{matrix} ^1W_{h\hat{x}} & ^1W_{hu} & ^2b \end{matrix} \cdot \begin{matrix} \hat{x}(k+1) \\ u(k+1) \\ 1 \end{matrix} \right] \\ &= ^2W_h \cdot g(^1W_{h\hat{x}} \cdot \hat{x}(k+1) + ^1W_{hu}u(k+1) + ^2b) \\ &= ^2W_h \cdot g(^1W_{h\hat{x}} \cdot ^2W_f \cdot ^1O(k+1) + ^1W_{hu}u(k+1) + ^2b). \end{aligned} \tag{6}$$

再令 $^3W = ^2W_h, ^2W = ^1W_{h\hat{x}} \cdot ^2W_f, ^4W = ^1W_{hu}, ^1W = ^1W_{fu}$, 则有

$$y(k+1) = {}^3W \cdot {}^2O(k+1), \quad (7a)$$

$${}^2O(k+1) = g({}^2W \cdot {}^1O(k+1) + {}^4W \cdot u(k+1) + {}^2b), \quad (7b)$$

$${}^1O(k+1) = g({}^H W \cdot O(k) + {}^1W \cdot u(k) + {}^1b). \quad (7c)$$

又由文献[2]中定理1, 存在相应维数的权值矩阵 W_f^* , W_h^* 使得(3)式成立, 则由上述推导和定义, 存在 ${}^1W^*$, ${}^2W^*$, ${}^3W^*$, ${}^H W^*$, ${}^4W^*$, ${}^1b^*$, ${}^2b^*$, 使(3)式成立.

3 网络的学习算法

设 Θ 表示神经网络中所有待求的权及阈值, 为简单计, 只对 MISO 系统推导, MIMO 系统类似. 令预报误差 $\varepsilon(k) = y(k) - \hat{y}(k)$, 定义目标函数

$$J(\Theta) = \frac{1}{2N} \sum_{i=k-N+1}^k \varepsilon^2(i, \Theta), \quad (8)$$

N 是数据长度. 则有

$$\Theta(k) = \Theta(k-1) - \eta \nabla J(\Theta(k-1)),$$

$$\nabla J(\Theta) = -\frac{1}{N} \sum_{i=k-N+1}^k \Psi(i, \Theta) \varepsilon(i, \Theta).$$

式中 η 为迭代步长, 取 $N=1$. $\Psi(i, \Theta)$ 为 \hat{y} 对各权值的导数.

把动态神经网络写为和式

$$\hat{y}(k) = \sum_{i=1}^{n_2} {}^3W_{1i} {}^2O_i(k),$$

$${}^2O_i(k) = g\left(\sum_{j=1}^{n_1} {}^2W_{ij} {}^1O_j(k) + \sum_{l_2=1}^r {}^4W_{il_2} u_{l_2}(k) + {}^2b_i\right),$$

$${}^1O_j(k) = g\left(\sum_{l_1=1}^{n_1} {}^H W_{jl_1} {}^1O_{l_1}(k-1) + \sum_{l_2=1}^r {}^1W_{jl_2} u_{l_2}(k-1) + {}^1b_j\right).$$

注意到隐层的动态性, 对 Θ 中的不同分量, 有

$$\frac{d\hat{y}(k)}{d{}^3W_{1i}} = {}^2O_i(k),$$

$$\frac{d\hat{y}(k)}{d{}^2W_{ij}} = {}^3W_{1i} \cdot {}^2O_i(k) (1 - {}^2O_i(k)) \cdot {}^1O_j(k),$$

$$\frac{d\hat{y}(k)}{d{}^4W_{il_2}} = {}^3W_{1i} \cdot {}^2O_i(k) (1 - {}^2O_i(k)) u_{l_2}(k),$$

$$\frac{d\hat{y}(k)}{d{}^2b_i} = {}^3W_{1i} \cdot {}^2O_i(k) (1 - {}^2O_i(k)) \triangleq z_i(k),$$

$$\begin{aligned} \frac{d\hat{y}(k)}{d{}^H W_{jl_1}} &= \sum_{i=1}^{n_2} z_i(k) \cdot {}^2W_{ij} \cdot \frac{d{}^1O_j(k)}{d{}^H W_{jl_1}} \\ &= \sum_{i=1}^{n_2} z_i(k) \cdot {}^2W_{ij} \cdot {}^1O_j(k) \cdot (1 - {}^1O_j(k)) ({}^1O_{l_1}(k-1) \\ &\quad + {}^H W_{jj} \frac{d{}^1O_j(k-1)}{d{}^H W_{jl_1}}), \end{aligned}$$

$$\frac{d\hat{y}(k)}{d{}^1W_{jl_2}} = \sum_{i=1}^{n_2} z_i(k) \cdot {}^2W_{ij} \cdot \frac{d{}^1O_j(k)}{d{}^1W_{jl_2}},$$

$$\frac{d\hat{y}(k)}{d{}^1b_j} = \sum_{i=1}^{n_2} z_i(k) \cdot {}^2W_{ij} \cdot \frac{d{}^1O_j(k)}{d{}^1b_j}.$$

其中

$$\frac{d^1 O_j(k)}{d^1 W_{j_{l_1}}} = {}^1 O_{l_1}(k) \cdot (1 - {}^1 O_j(k)) ({}^1 O_j(k-1) + {}^H W_{j_{l_1}} \frac{d^1 O_j(k-1)}{d^1 W_{j_{l_1}}}),$$

$$\frac{d^1 O_j(k)}{d^1 W_{j_{l_2}}} = {}^1 O_j(k) \cdot (1 - {}^1 O_j(k)) (u_{l_2}(k) + {}^H W_{jj} \frac{d^1 O_j(k-1)}{d^1 W_{j_{l_2}}}),$$

$$\frac{d^1 O_j(k)}{d^1 b_j} = {}^1 O_j(k) \cdot (1 - {}^1 O_j(k)) (1 + {}^H W_{jj} \frac{d^1 O_j(k-1)}{d^1 b_j}),$$

$$i = 1, \dots, n_2; j = 1, \dots, n_1; l_1 = 1, \dots, n_1; l_2 = 1, \dots, r.$$

因 $\Psi(k, \Theta)$ 是动态的所以又叫动态 BP 法。

4 仿真与分析

例.

$$\begin{cases} x_1(k+1) = \frac{x_1(k)}{1+x_2^2(k)} + u_1(k), \\ x_2(k+1) = \frac{x_1(k)x_2(k)}{1+x_2^2(k)} + u_2(k), \\ y_p(k) = x_2(k). \end{cases}$$

分别用 RFNN 的串并法和本文 RHNN 的并联法对此系统的输入输出关系辨识, 若事先已知 y_p 与 $u_{1,2}$ 无直接关系, 可令 $W=0$. 仿真参数及结果如表 1、图 2、图 3 所示。

表 1

	RFNN(串并法辨识)	RHNN(并联辨识)
算法	BP 法, $\eta=0.05$	动态 BP 法, $\eta=0.25$
网络结构	4-45-1	2-10-8-1
训练输入	$u_{1,2} = \sin \frac{2k\pi}{25}, y_p(k), y_p(k-1)$	$u_{1,2} = \sin \frac{2k\pi}{25}$
训练步数	$k=20,000$	$k=20,000$
预测输入	$u_{1,2} = \sin \frac{2k\pi}{25}, \hat{y}(k), \hat{y}(k-1)$	$u_{1,2} = \sin \frac{2k\pi}{25}$
预测结果	图 2	图 3
结论	1)	2)

由上例及其它大量仿真可得如下结论:

1) RFNN 对满足状态方程的非线性系统的输入输出关系的全局逼近能力不强。

2) RHNN 对这种输入输出关系的描述能力比较强。且此关系的逼近精度与训练输入的频带有关, 若预测输入在训练输入频带附近则预测精度高。

本文提出用含动态隐层的四层前馈网对非线性系统进行辨识, 证明了此方案的可行性。网络学习算法为动态 BP 法。仿真表明此种结构的网络通过学习对满足状态方程的非线性系统的输入输出关系能进行较好的辨识。

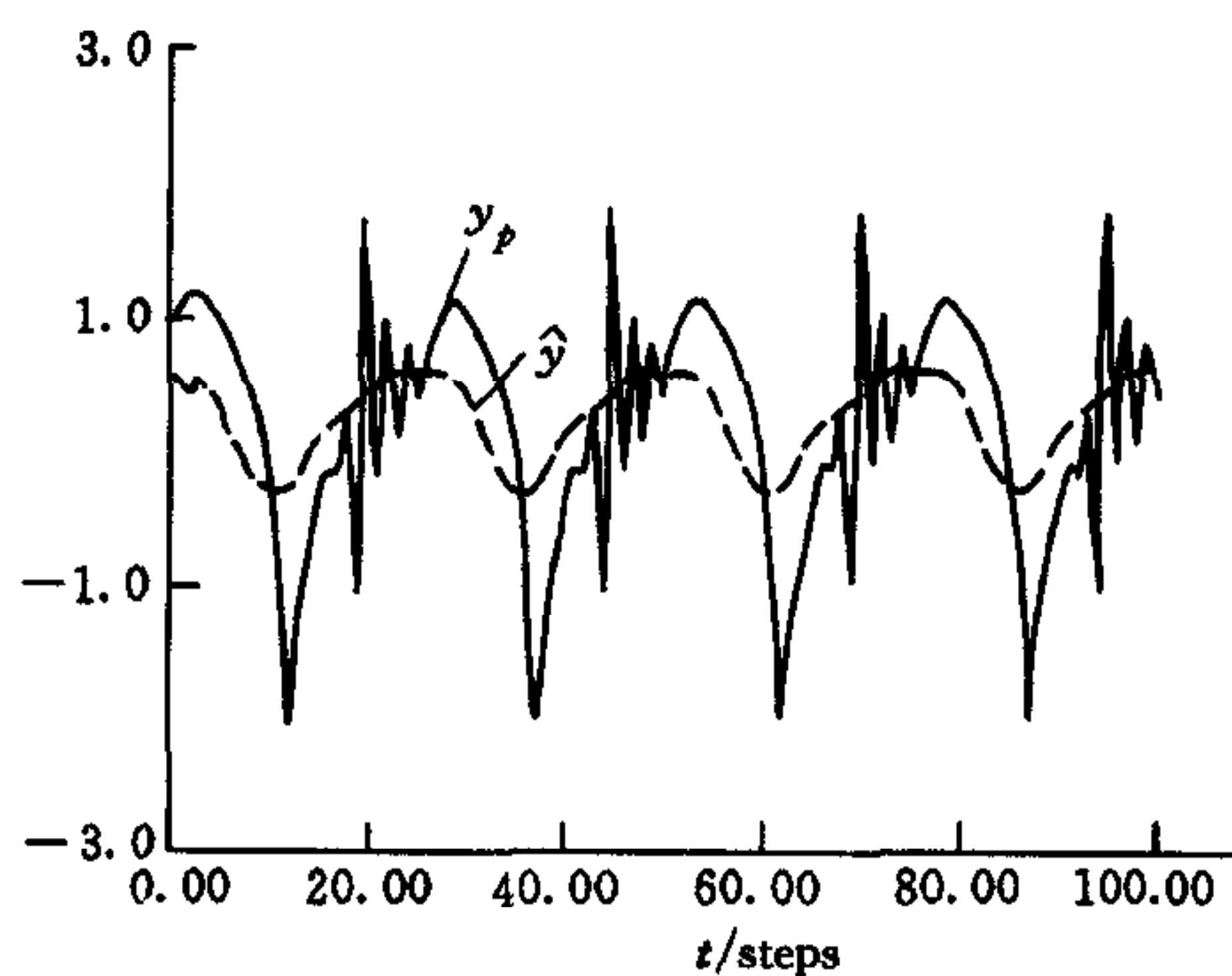


图2 用 RFNN 的预测结果

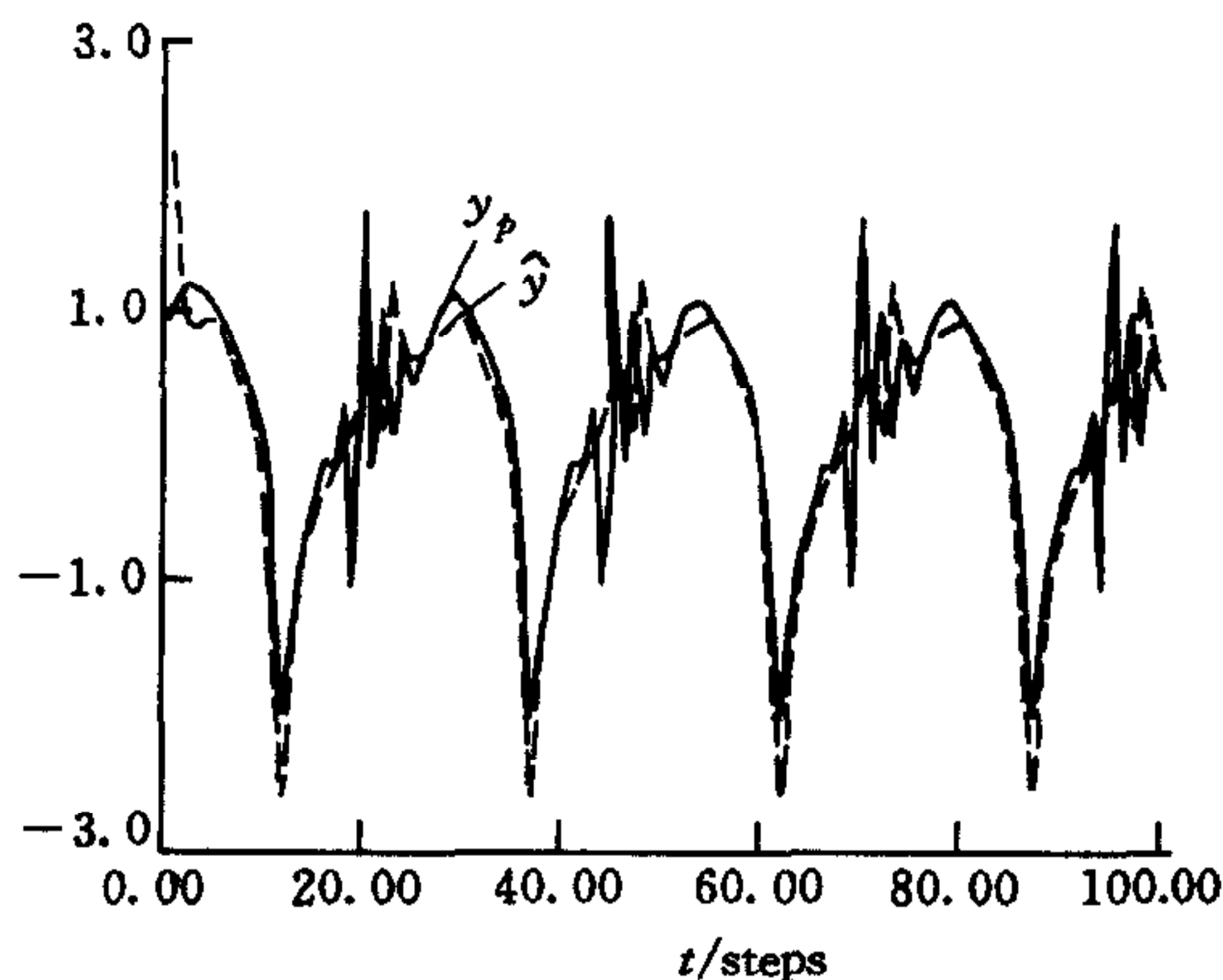


图3 用 RHNN 的预测结果

参 考 文 献

- [1] Leontaritis I J, Billings S A. Input-output parametric models for nonlinear systems, part I; Deterministic non-linear systems. *Int. J. Control*, 1985, 41(2):303—328.
- [2] 刘延年, 冯纯伯. 用神经网络进行非线性离散动态系统辨识的可行性. *控制理论与应用*, 1994, 11(4):413—420.

NONLINEAR SYSTEM IDENTIFICATION USING FEEDFORWARD NEURAL NETWORKS WITH RECURRENT HIDDEN LAYER

BAO XIAOHONG JIA YINGMIN

(The Seventh Research Division, Beijing University of Aeronautics and Astronautics, Beijing 100083)

Abstract This paper presents a feedforward neural network which contains a dynamic hidden layer (RHNN) for the modeling of a nonlinear system and the feasibility of this approach is proved. The learning algorithm is the dynamic BP. In the simulation a comparison between RHNN and RFNN for modeling a nonlinear system is made.

Key words Feedforward neural networks, system identification, dynamic BP algorithm.