

用神经网络建立非线性系统模型研究*

杨 熔 李永华 苏义鑫

(武汉工学院电子电气工程系, 430072)

摘要: 本文针对多层神经网络结构, 运用递推预报误差(RPE)算法对离散非线性系统进行辨识研究. 作为应用实例, 本文对一个工业实际对象进行了神经网络动态建模. 研究表明, 神经网络方法是用于带有非线性特性工业过程建模的有效方法.

关键词: 多层神经网络; 辨识; 递推预报误差法; 非线性系统; 建模; PWM 直流调速系统

1 引言

神经网络的出现为一类输入输出关系呈高度非线性的系统提供了建模的有效工具. 因此, 探讨和研究神经网络建模的算法、方法和性质显得尤为重要. 这方面已有一些新的进展, 如英国的 Billings 等人将一种传统的参数估计方法即预报误差法(RPE)运用于多层神经网络的学习训练, 以取代反向传播(BP)算法. 该算法体现出收敛速度快、预报精度高等优点^[1].

本文进一步运用 RPE 算法对实际工业过程进行辨识研究. 首先介绍用神经网络建立非线性模型机理及单隐层神经网络的 RPE 算法; 在此基础上研究一个带有非线性特性的对象——PWM 直流调速系统的神经网络动态建模, 这是将神经网络方法用于快速的非线性系统辨识的尝试. 文章对所叙述对象的非线性特性进行了分析, 对神经网络的结构、输入和隐层节点数的选取进行了仿真研究. 研究结果充分表明: 神经网络方法是用于复杂工业过程建模的有效方法; 它较之传统的线性系统辨识方法, 更能自然地揭示和逼近过程的输入输出关系.

2 神经网络建模机理

2.1 非线性系统描述

一类广泛的离散时间非线性系统可以用外生输入非线性自回归滑动平均模型(NARMAX)来表示. 该模型根据自变量是输入、输出及噪声的延迟信号的非线性函数描述非线性过程. 一个单输入单输出非线性系统的近似 NARMAX 模型可用下式表示

$$y(t) = f(y(t-1), \dots, y(t-n_y), u(t-1), \dots, u(t-n_u)) + e(t). \quad (1)$$

式中 $y(t)$, $u(t)$ 分别是输出、输入; n_y , n_u 分别是输出、输入的最大延迟; $e(t)$ 是不相关噪声; $f(\cdot)$ 是非线性函数. 这种非线性关系一般非常复杂, 难以用一类统一模型表达. 神经网络的出现为这类问题的建模带来了希冀.

2.2 多层神经网络

多层神经网络由输入层、隐层及输出层组成. 整个网络的特性决定于相邻层间神经元

* 湖北省自然科学基金资助项目.

本文于 1993 年 9 月 11 日收到, 1993 年 12 月 27 日收到修改稿.

的联结权值以及隐层中神经元的阈值. 一个描述单输入单输出非线性系统的单隐层神经网络结构图如图 1 示. n_0, n_1, n_2 分别是输入层、隐层及输出层中神经元的数目. $x_i (i = 1, \dots, n_0)$ 及 $\hat{y}$ 是网络的输入输出值. $x_i^l, b_i^l (i = 1, \dots, n_1)$ 是隐层中第 i 个神经元的输出值和阈值. w_{ij}^k 表示第 $k-1$ 层中第 j 个神经元对第 k 层中第 i 个神经元的联结权值. 取隐层节点的作用函数为 S 型函数,

$$g(z) = \frac{1}{1 + e^{-z}} \quad (2)$$

则整个网络的输入输出关系为

$$x_j^1 = g\left(\sum_{i=1}^{n_0} w_{ij}^1 x_i + b_j^1\right), \quad (3)$$

$$\hat{y} = \sum_{j=1}^{n_1} w_j^2 x_j^1. \quad (4)$$

2.3 用神经网络建模

设 $\theta = [\theta_1, \dots, \theta_n]^T$ 表示神经网络中所有待求的权及阈值, 且 $n = (n_0 + 1) \times n_1 + n_1$ 是待求参数的数目. 由输入输出的延迟项构成输入向量. $\varepsilon(t)$ 为实际过程输出与神经网络预报输出 $\hat{y}$ 之差. 通过某种训练算法使 θ 在某种准则下得以确定, 用神经网络建立非线性过程模型原理图见图 2.

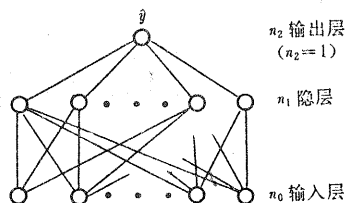


图 1 单隐层神经网络

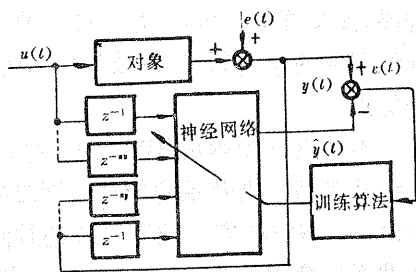


图 2 用神经网络建模示意图

3 神经网络的 RPE 算法

3.1 RPE 算法基本原理

通过极小化预报误差准则来获得参数估计的方法叫预报误差法. 准则函数 J 定义为:

$$J(\theta) = \frac{1}{2N} \sum_{t=1}^N \varepsilon^T(t, \theta) \varepsilon(t, \theta). \quad (5)$$

式中 $\varepsilon(t)$ 是预报误差向量, N 是数据长度. RPE 算法的基本原理是沿 J 的 Gauss-Newton 搜索方向修正未知参数向量, 从而使 $J \rightarrow \min$. 参数向量的修正算式为

$$\theta(t) = \theta(t-1) + s(t) \mu \theta(t-1).$$

$\mu(\theta)$ 是 Gauss-Newton 搜索方向. 它定义为

$$\mu(\theta) = -[H(\theta)]^{-1} \nabla J(\theta).$$

式中 $\nabla J(\theta)$ 是 $J(\theta)$ 关于 θ 的梯度, $H(\theta)$ 是 $J(\theta)$ 的 Hessian 矩阵; 它们分别是 $J(\theta)$ 关于 θ 的一阶和二阶微分. 且有

$$\nabla J(\theta) = -\frac{1}{N} \sum_{t=1}^N \Psi(t, \theta) \varepsilon(t, \theta), \quad H(\theta) = \sum_{t=1}^N \Psi(t, \theta) \Psi^T(t, \theta).$$

其中 $\Psi(t, \theta) = \left[\frac{d\hat{y}(t, \theta)}{d\theta} \right]^T$ 是网络的一步预报值对 θ 的一阶微分. 基于上述原理的递推预

报误差算法是:

$$\begin{cases} \varepsilon(t) = y(t) - \hat{y}(t), \\ P(t) = \frac{1}{\lambda(t)} \{P(t-1) - P(t-1)\Psi(t)[\lambda(t) + \Psi^T(t)P(t-1)\Psi(t)]^{-1}\Psi^T(t)P(t-1)\}, \\ \hat{\theta}(t) = \hat{\theta}(t-1) + P(t)\Psi(t)\varepsilon(t), \end{cases} \quad (6)$$

$\lambda(t)$ ——遗忘因子.

迭代初期置 $\lambda(t) < 1$, 从而实现自适应算法; 而当 $t \rightarrow \infty$ 时, 取 $\lambda \rightarrow 1$, 按下式可以达到这个目的.

$$\lambda(t) = \lambda_0 \lambda(t-1) + (1 - \lambda_0).$$

这里 λ_0 和初始遗忘因子 $\lambda(0)$ 根据需要设置, 本文取 $\lambda_0 = 0.99, \lambda(0) = 0.95$.

3.2 RPE 算法用于神经网络的具体实现

对于图 1 所示的单隐层神经网络, 利用(4)式不难得到 Ψ 阵($n \times 1$)的元素:

$$\Psi = \frac{d\hat{y}}{d\theta_i} = \begin{cases} x_k^1, & \text{当 } \theta_i = w_k^2, \quad 1 \leq k \leq n_1, \\ x_k^1(1 - x_k^1)w_k^2, & \text{当 } \theta_i = b_k^1, \quad 1 \leq k \leq n_1, \\ x_k^1(1 - x_k^1)w_k^2x_m, & \text{当 } \theta_i = w_{km}^1, \quad 1 \leq k \leq n_1, 1 \leq m \leq n_0. \end{cases} \quad (7)$$

至此可归纳出 RPE 算法用于单隐层神经网络的具体流程:

① 取权和阈值为较小随机值; 取 $P(0)$ 为对角阵, 对角线上的元素值在 10^4 左右; 取 λ 为合适值.

② 根据网络输入, 按式(3), (4) 分别计算隐层节点 x_k^1 及输出节点 $\hat{y}$ 的值.

③ 按(7)式构成 Ψ 阵.

④ 按(6)式求预报误差 $\varepsilon(t)$, P 阵及参数序列 $\hat{\theta}(t)$. ②至④步反复进行, 直至收敛.

下面对一个对象进行神经网络 RPE 建模仿真研究. 对象模型是

$$y(t) = \frac{0.6}{1 + \exp\{-(0.5u(t-1) + 0.4y(t-1) + 0.1)\}}.$$

显然可以建立一个 $n_0 = 2, n_1 = n_2 = 1$ 的神经网络模型. 输入信号 $u(t)$ 为六级 PRBS 序列.

输入向量 $X(t) = [u(t-1) \quad y(t-1)]^T$,

参数向量 $\theta(t) = [\theta_1 \quad \theta_2 \quad \theta_3 \quad \theta_4]^T = [w_{11}^1 \quad w_{12}^1 \quad b_1^1 \quad w_1^2]^T$,

Ψ 阵元素 $\Psi = [\varphi_1 \quad \varphi_2 \quad \varphi_3 \quad \varphi_4]^T = [a_1x_1 \quad a_1x_2 \quad a_1 \quad x_1^1]^T$, $a_1 = x_1^1(1 - x_1^1)w_1^2$.

递推 1000 次后得参数为

$$\theta = [0.4999 \quad 0.3999 \quad 0.09985 \quad 0.6000]^T.$$

可见结果是非常理想的. 而对该例施行 BP 算法, 由于 $\eta_w, \eta_b, \alpha_w, \alpha_b$ 等参数难以折衷处理, 始终得不到理想结果.

4 动态建模实例

一般设计由直流脉宽调速系统(PWM)加位置环构成的位置伺服系统时,没有考虑 PWM 系统中存在的非线性因素. 双环 PWM 调速系统的线性动态结构图如图 3 示. 而实际上因速度环的特性随给

定工况、电源电压、负载和转动惯量而变化,从而使系统具有时变特性^[3]. 此外,系统中库仑摩擦的影响;误差信号限幅的影响;齿隙的影响以及调节器和脉宽调制放大器的饱和非线性的影响均是影响系统性能的非线性因素.

文献[3]对某厂位置伺服系统中的 PWM 进行了现场辨识实验,建立了速度环的线性动态模型. 为了进一步提高位置伺服系统的设计水平,为高性能的自适应控制打下基础,弄清 PWM 系统非线性特性是很重要的. 在此考虑用神经网络方法对 PWM 系统进行动态建模. 关于辨识实验装置及试验信号参数选择的详细描述参见[3]. 整个速度环作为辨识对象,其输入信号是在速度给定信号 $n_{\text{给}} = 2\text{V}$ 之上迭加幅值为 0.2V 的 PRBS 信号;输出信号是测速电机输出电压.

用来建立 PWM 系统神经网络的结构确定为 $n_0 = 3, n_1 = 2$. 输入向量定义为

$$X(t) = [u(t-1)u(t-2)y(t-1)]^T.$$

图 4 示出实测试验数据及用 RPE 算法训练 900 次后神经网络预报输出和残差序列. 图 5 示出神经网络的预报输出

与实测阶跃响应的拟合曲线. 由此可见,神经网络能满意地建立 PWM 系统的动态模型.

引入误差指数 EI 作为表征模型拟合精度的指标. EI 定义为:

$$\text{EI} = \sqrt{\sum \varepsilon^2(t) / \sum y^2(t)}.$$

其中 $\varepsilon(t) = \hat{y}(t, \hat{\theta}) - y(t)$. $\hat{\theta}$ 是递推 900 次后所得的参数向量. 改变隐层节点数目后 EI 值的比较见表 1. 改变输入节点及输入向量分配后所得的 EI 值的比较见表 2.

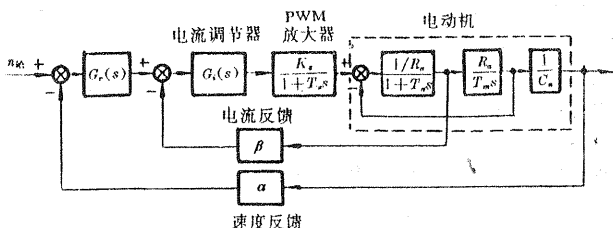


图 3 双环 PWM 系统结构框图

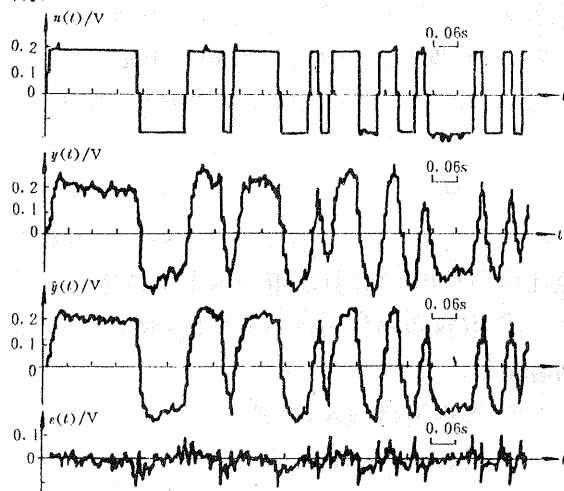


图 4 实测数据及辨识结果

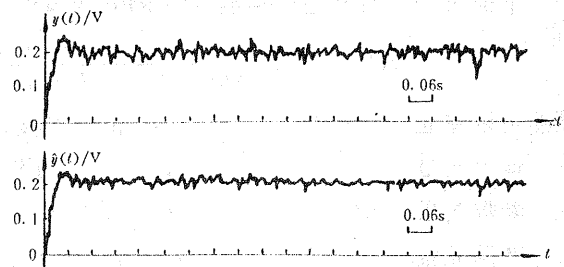


图 5 阶跃响应拟合结果

表 1 改变隐层节点数目后的 EI 值比较

n_0	输入向量 $X(t)$	n_1	EI
3	$[u(t-1)u(t-2)y(t-1)]^T$	1	0.6258
		2	0.1907
		3	0.1905
		4	0.1923
		5	0.1961

表 2 改变输入节点及输入向量分配的 EI 值比较

n_0	输入向量 $X(t)$	n_1	EI
1	$[u(t-1)]^T$	2	0.6964
2	$[u(t-1)y(t-1)]^T$		0.1946
3	$[u(t-1)u(t-2)y(t-1)]^T$		0.1907
	$[u(t-1)y(t-1)y(t-2)]^T$		0.1892
	$[u(t-1)u(t-2)u(t-3)y(t-1)]^T$		0.1894
4	$[u(t-1)u(t-2)y(t-1)y(t-2)]^T$	2	0.1883
	$[u(t-1)y(t-1)y(t-2)y(t-3)]^T$		0.1857
5	$[u(t-1)u(t-2)y(t-1)y(t-2)y(t-3)]^T$		0.1865

由此可见,本文神经网络模型结构是合理的. 另外还可看出,输入节点数 n_0 及输入向量的合理构成对拟合精度的影响是很大的. 结合本例,增加输入节点数目,增加输入向量中 $y(t)$ 的延迟项,将有助于提高建模精度.

5 结束语

本文的主要结论是:

- 1) 神经网络具有学习非线性特性的非凡能力;PWM 系统用神经网络建模,弥补了原线性理论分析的不足.
- 2) RPE 算法具有比 BP 算法更好的收敛性、稳定性,具有广泛的使用价值.
- 3) 多层神经网络是建模的主要模型;对单隐层网络而言,在合理选择隐层节点数的基础上,适当增加输入节点数目及合理构成输入向量对建模精度有很大影响.

参 考 文 献

[1] Billings, et al. A Comparison of the Backpropagation and Recursive Prediction Error Algorithms for Training Neural Networks. Mechanical systems and Signal processing, 1991, 233—255

[2] 焦李成. 神经网络系统理论. 西安:西安电子科技大学出版社,1990

[3] 徐枋同,杨熔等. 双环直流脉宽调速系统动态模型辨识. 电气传动, 1993, 23(2):36—41

Research on Modelling Non-Linear Systems Using Neural Networks

YANG Rong, LI Yonghua and SU Yixin

(Department of Electronic & Electric Engineering, Wuhan Institute of Technology • Wuhan, 430072, PRC)

Abstract: This paper discusses the use of recursive prediction error algorithm of training multilayered neural networks for discrete time non-linear system identification. Results based on the simulated system and the dynamic modelling of an industrial example indicate that the recursive prediction error algorithm is attractive and effective for modelling of complicated industrial processes which possess non-linear characteristics.

Key words: multilayered neural networks; identification; recursive prediction error algorithm; non-linear system; model; PWM DC speed control system

本文作者简介

杨 熔 女. 1958 年生. 1982 年毕业于武汉工学院工业自动化专业, 获学士学位. 尔后留校任教. 1990 年获武汉工学院工业自动化专业硕士学位. 现任武汉工学院自动化教研室讲师. 发表论文十余篇. 目前主要研究方向: 神经网络在系统辨识及自适应控制中的应用.

李永华 女. 1938 年生. 武汉工学院教授, 电子电气工程系自动控制研究所所长, 中国汽车工程学会“机器人”分会秘书长, 湖北、武汉自动化学会理事. 1960 年华中理工大学工业电气自动化专业毕业. 1982 年至 1984 年在东京工业大学制御工学学科作访问学者. 1988 年曾为东京工业大学客座研究员. 在国内外发表论文 30 余篇. 主要研究方向为计算机数控(CNC). 控制系统的分析和综合, 工程过程的建模与仿真等.

苏义鑫 1965 年生. 1985 年毕业于武汉水利电力大学工业自动化专业, 获学士学位. 1988 年获东南大学自动控制理论及应用专业硕士学位. 现任武汉工学院自动化教研室讲师. 发表论文若干篇. 目前研究兴趣: 神经网络在系统辨识及自适应控制中的应用.