

基于动态 BP 算法的非线性滞后系统辨识

韩冰^{*1,2}, 韩敏¹

(1. 大连理工大学 控制科学与工程学院, 辽宁 大连 116024;

2. 上海船舶运输科学研究所, 上海 200135)

摘要: 非线性动态系统的建模一直是控制领域的重要问题之一. 针对这一问题, 特别是包含滞后环节的非线性系统建模问题, 提出了一种引入自适应延迟的动态 BP(back propagation) 学习算法. 该算法在传统多层感知机神经网络结构基础上, 在网络的第 1 隐层和输出层分别引入可调节的自适应延迟参数, 通过误差梯度对其进行修正, 实现了对延迟参数的辨识. 仿真结果表明, 所提出的方法能够有效实现对非线性滞后系统的辨识, 并能够对系统的延迟时间进行准确估计.

关键词: 非线性系统; 时滞; 系统辨识; 神经网络

中图分类号: TP183 **文献标志码:** A

0 引言

系统辨识作为控制科学的核心问题长期以来一直受到广泛的关注. 非线性系统, 特别是包含滞后的非线性系统辨识一直是控制领域的难题之一^[1]. 对非线性滞后系统的辨识, 以往传统的方法是从输入输出数据入手, 采用 Hemmserstein 模型^[2]、Wiener 模型^[3]等模型结构进行建模. 但对于结构和参数未知的系统, 以上方法很难对实际系统进行表达. 随着智能建模理论的发展, 局部模型近似^[4]、模糊模型^[5]、支持向量机^[6]以及遗传算法^[7]等得到了广泛的采用, 但这些方法对先验知识仍具有较强的依赖性. 因此, 如何对模型参数完全未知非线性滞后系统进行辨识, 如何辨识得到该系统的延迟时间成为非线性滞后系统建模研究的关键问题.

对模型未知的滞后系统延迟时间辨识问题, Balestrino 等^[8]、陆燕等^[9]提出了基于感知机神经网络进行延迟参数辨识的方法, 但这些方法易受到网络参数和辨识对象等因素的影响, 在实际应用中存在一定的局限. Ren 等^[10]扩展了延迟自由神经网络的思想, 提出了可对一类滞后系统进行

有效延迟参数估计的神经网络算法. Han 等^[11]在分多支延迟神经网络中引入可调的延迟参数, 对包含滞后的非线性对象进行建模, 并应用于预测控制中. 这些研究成果表明, 通过神经网络对模型未知的非线性滞后系统辨识是较为可行的方法.

本文利用神经网络具有很好的函数逼近能力, 结合多层感知机网络的特点, 分别在多层感知机网络的第 1 隐层和输出层引入带有自适应延迟参数的动态神经元, 提出一种包含动态特性的动态 BP 学习算法. 该算法对于黑箱非线性滞后系统, 能够通过第 1 隐层延迟参数来自适应选择模型输入变量, 避免采用非线性含输入自回归(NARX)等建模方法进行输入变量选择时所需要的复杂计算^[12].

1 动态 BP 算法

为了改善传统 BP 算法不具备动态特性的缺陷, 本文在传统感知机神经网络的第 1 隐层和输出层引入动态神经元, 构成新的动态 BP 神经网络. 动态 BP 神经网络的结构如图 1 所示. 其结构与传统 BP 神经网络相似, 但在第 1 隐层和输出层引入了和网络权值对应的延迟参数.

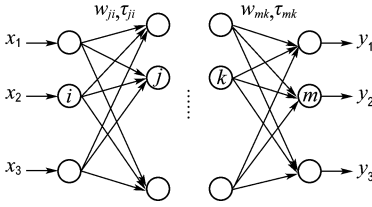


图 1 动态 BP 神经网络结构图

Fig. 1 Basic structure of dynamic BP neural network

第 1 隐层和输出层的动态神经元输入如下式所示:

$$net_j(t) = \sum_{i=1}^N w_{ji} o_i(t - \tau_{ji}) \quad (1)$$

式中: $net_j(t)$ 表示第 j 个动态神经元的输入; w_{ji} 表示前一层第 i 个神经元到该层第 j 个神经元之间的权值; o_i 表示该神经元的第 i 个输入, 共有 N 个输入; τ_{ji} 表示前一层第 i 个神经元到该层第 j 个神经元之间的延迟参数。

动态神经元的输出为

$$y(t) = f(net_j(t)) = f\left(\sum_{i=1}^N w_{ji} o_i(t - \tau_{ji})\right) \quad (2)$$

这里 f 为神经元激活函数, 本文神经元激活函数均选取 Sigmoid 函数, 如下式所示:

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (3)$$

网络学习过程中, 网络的权值可按传统 BP 算法^[13]进行更新. 网络误差函数为

$$E(t) = \frac{1}{2} \sum_{k=1}^K (y_k(t) - o_k(t))^2 \quad (4)$$

第 1 隐层和输出层的延迟参数可根据网络学习误差的梯度进行实时调整. 对输出层的延迟参数, 网络误差函数的梯度下降方法可给出延迟参数更新值为

$$\Delta\tau_{mk} = -\eta_o \left(\frac{\partial E(t)}{\partial o_k(t)} \right) \left(\frac{\partial o_k(t)}{\partial \tau_{mk}} \right) = \eta_o \delta_m(t) \frac{\partial o_m(t)}{\partial \tau_{mk}} \quad (5)$$

其中 $\Delta\tau_{mk}$ 为网络输出层第 m 个神经元与之前一层的第 k 个神经元之间延迟参数 τ_{mk} 的修正量, η_o 为网络输出层延迟参数的学习率, $o_m(t)$ 为网络输出层第 m 个节点的输出, $\delta_m(t)$ 为 t 时刻网络的学习误差, 即

$$\delta_m(t) = y(t) - o_m(t) \quad (6)$$

$\frac{\partial o_m(t)}{\partial \tau_{mk}}$ 可以按以下公式进行求解:

$$\frac{\partial o_m(t)}{\partial \tau_{mk}} = \frac{\partial o_m(t)}{\partial net_m(t)} \frac{\partial net_m(t)}{\partial \tau_{mk}} =$$

$$f'(net_m(t)) \tau_{mk} \frac{\partial o_k(t - \tau_{mk})}{\partial \tau_{mk}} \quad (7)$$

其中

$$\frac{\partial o_k(t - \tau_{mk})}{\partial \tau_{mk}} \approx \frac{o_k(t - \tau_{mk}) - o_k(t - \tau_{mk} - T)}{T}$$

f 为神经元激活函数, T 为系统采样时间.

对第 1 隐层延迟参数, 可按式(7)进行迭代更新:

$$\Delta\tau_{ji} = -\eta_i \left(\frac{\partial E(t)}{\partial o_j(t)} \right) \left(\frac{\partial o_j(t)}{\partial \tau_{ji}} \right) = \eta_i \delta_j^l(t) \frac{\partial o_j(t)}{\partial \tau_{ji}} \quad (8)$$

式中: $\Delta\tau_{ji}$ 为网络输入层第 i 个神经元与第 1 隐层的第 j 个神经元之间延迟参数的修正量, η_i 为网络第 1 隐层延迟参数的学习率, $o_j(t)$ 为网络第 1 隐层第 j 个节点的输出, 其中

$$\begin{aligned} \delta_j^l(t) &= \sum_{n=1}^{N^{l+1}} \delta_n^{l+1}(t) f'(net_n^{l+1}(t)) \tau_{nj}^{l+1}, \\ \frac{\partial o_j(t)}{\partial \tau_{ji}} &= f'(net_j^l(t)) w_{ji} \frac{\partial o_i(t - \tau_{ji})}{\partial \tau_{ji}}, \\ \frac{\partial o_i(t - \tau_{ji})}{\partial \tau_{ji}} &\approx \frac{o_i(t - \tau_{ji}) - o_i(t - \tau_{ji} - T)}{T} \end{aligned} \quad (9)$$

式中: $\delta_j^l(t)$ 表示 l 层 t 时刻误差函数, 第 1 隐层误差函数可由式(9)通过输出层误差递推得到. N^{l+1} 表示网络第 $l+1$ 层的节点数; $net_n^{l+1}(t)$ 表示 $l+1$ 层网络第 n 个节点 t 时刻的输入; $o_i(t)$ 表示神经网络的输入.

2 基于动态 BP 算法的系统辨识

一般的非线性系统可以表示为

$$\begin{aligned} y(t) &= f[y(t-1), y(t-2), \dots, y(t-N), \\ &\quad x(t-1), x(t-2), \dots, x(t-M)] \end{aligned} \quad (10)$$

其中 N, M 是最大的延迟范围. 利用动态 BP 算法对系统(10)进行建模, 使神经网络逼近系统(10)中的非线性函数 $f[\cdot]$. 由动态神经元输入式(1)可以将第 1 隐层的第 j 个神经元输出表示为

$$o_j^2(t) = f\left(\sum_{i=1}^{N^1} w_{ji} x(t - 1 - \tau_{ji})\right) \quad (11)$$

其中 N_1 为网络第 1 层神经元个数, $x(t - 1 - \tau_{ji})$ 为神经元的输入. 将第 2 层的输出式(11)代入到第 3 层的计算中有

$$o_j^3(t) = f\left[\sum_{j=1}^{N^2} w_{kj} f\left(\sum_{i=1}^{N^1} w_{ji} x(t - 1 - \tau_{ji})\right)\right]$$

其中延迟参数 τ_{ji} 通过使网络误差函数

$$E(t) = \frac{1}{2} \sum_{k=1}^K (y_k(t) - o_k(t))^2$$

最小化进行修正,这样经过层层迭代后,网络的输出可以表示为

$$o_k^L(t) = N[x(t-1), x(t-2), \cdots, x(t-1-2\tau_{\max})] \tag{12}$$

其中 L 表示网络输出层, $N[\cdot]$ 为神经网络所建立的非线性函数, $\tau_{\max}$ 是所有延迟参数在训练结束后的最大值. 训练过程中将前一时刻的输出 $y(t-1)$ 作为输入,则网络的输出可以表示为

$$o_j^L(t) = \sigma[y(t-1), y(t-2), \cdots, y(t-N_\tau), x(t-1), x(t-2), \cdots, x(t-M_\tau)] \tag{13}$$

其中 N_τ, M_τ 是可能的最大延迟. 从以上表达式可以看出,引入动态神经元使多层感知机网络具有了动态特性,能够对一类包含滞后特性的非线性对象进行建模. 同时,从式(10)和(13)可以看出,通过辨识得到的神经网络延迟参数可以对滞后对象的延迟进行估计,从而为滞后对象的延迟时间辨识提供新方法.

3 仿真实例

利用本文提出的动态 BP 算法,分别对理论数学模型和实际生产过程进行系统辨识仿真,进而验证该算法的有效性.

3.1 大滞后模型辨识仿真实验

利用文献[14]中给出的动态模型来检验自适应算法的有效性. 模型的差分方程如下:

$$y(k+1) = \frac{A(k+1)}{B(k+1)} \tag{14}$$

式中

$$A(k+1) = y(k) \cdot y(k-1) \cdot y(k-2) \times u(k-1) \cdot [y(k-2) - 1] + u(k-10)$$

$$B(k+1) = 1 + y^2(k-1) + y^2(k-2)$$

从模型的方程来看,当前的输出由先前的输出和输入决定,相对输入有 10 个单位的延迟. 用于训练的教师信号由 1 000 组输入输出序列组成. 其中前 500 组用于训练网络,后 500 组用来检验网络泛化能力. 输入为

$$u(k) = \begin{cases} \sin(\pi k/25); & 0 \leq k < 250 \\ 0.8\sin(2\pi k/250) + 0.2\sin(2\pi k/25); & 250 \leq k < 750 \\ 0.3\sin(\pi k/25) + 0.1\sin(\pi k/32) + 0.6\sin(\pi k/10); & 750 \leq k < 1\,000 \end{cases}$$

由 $u(k)$ 产生的教师信号如图 2 所示.

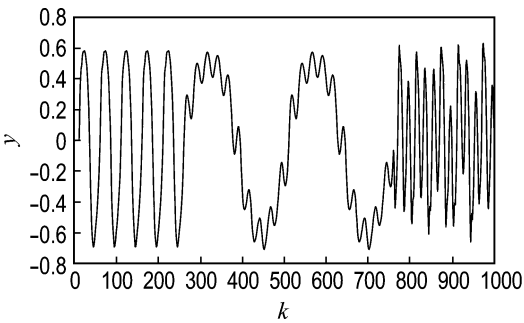


图 2 网络学习和测试信号

Fig. 2 Learning and test signal of the network

利用 C 语言编程实现本文所提出的动态 BP 学习算法对该模型的辨识,仿真条件如表 1 所示. 网络中共有 45 个延迟变量参数,第 1 隐层和输出层初始延迟值分别设置为 1 和 12. 经过 20 000 次学习之后网络的学习误差和泛化误差分别为 0.003 030、0.005 038,训练后网络的输出与教师信号的比较如图 3 所示. 从图 3 中可以看出,训练后的动态 BP 网络很好地体现了非线性系统的动态特性. 训练结束后,第 1 隐层延迟时间参数共 10 个值为 1,其余 20 个延迟参数为 0;输出层共 6 个延迟参数为 10,其余 9 个值为 9,可见输出层的延迟参数对辨识对象的纯滞后时间有较好的辨识能力.

表 1 仿真条件

Tab. 1 Simulation conditions

参数名称	参数值
网络结构	2-15-15-1
权值学习率	0.001
延迟学习率	0.001
节点函数	$f(x) = A \left(\frac{1 - e^{-x}}{1 + e^{-x}} \right)$
节点幅值 A	3
学习次数	20 000

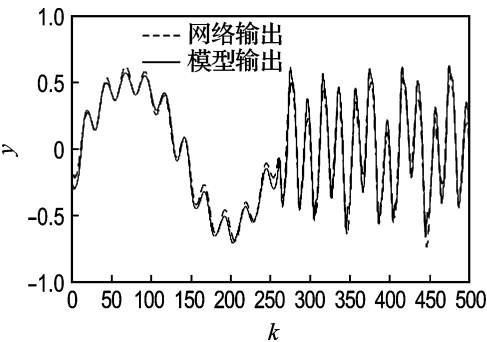


图 3 泛化信号同教师信号的比较

Fig. 3 Comparison between generalizing signal and teacher signal

为验证引入自适应延迟参数在算法学习过程中的作用,采用传统 BP 算法与本文提出的引入延迟参数的动态 BP 算法进行比较. 利用相同的网络结构采用传统 BP 算法经过 20 000 次学习后得到的学习误差和泛化误差分别为 0. 003 712、0. 009 019. 通过误差的比较可以看出,引入了自适应延迟参数加快了网络的收敛速度,并且网络模型的泛化能力有了较大的提高,网络具有更好的表达系统动态特性的能力.

3.2 大滞后模型辨识仿真实验实例验证

为验证本文所提出的动态 BP 网络对实际工业生产过程建模的有效性,选取某炼油厂连续重整装置的芳烃抽提塔为建模对象. 其中,溶剂比(环烃砜与进料量之比)作为可操作的控制量,抽提塔底温度为控制系统的输出变量. 现场采集用于训练神经网络的信号如图 4 所示. 其中虚线表示输入变量溶剂比 a 随时间的变化情况,其基值为 3. 25;实线表示抽提塔底温度 θ 的变化情况,其基值为 172 $^{\circ}\text{C}$. 系统采样时间为 2 min,共 100 组数据,前 90 组数据用于训练网络,后 10 组数据用于检验网络的泛化能力.

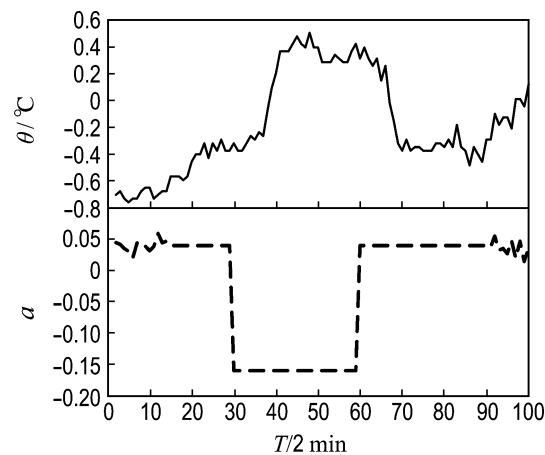


图 4 网络学习和测试信号(实例)

Fig. 4 Learning and text signal of the network (example)

采用本文所提出的自适应选择延迟时间的学习算法对该模型进行辨识,仿真条件如表 2 所示. 从图 4 可以看出,输出对阶跃响应的延迟在 10 个采样时间以内,因此将与输出节点相连的分支上的延迟时间初始值设为 10.

网络共 60 个延迟变量参数,第 1 隐层初始延迟值分别设置为 1. 经过自适应选择,第 1 隐层延迟时间参数共 10 个值为 1,其余延迟参数为 0;输出层共 12 个延迟参数为 8,其余 8 个为 9,因此可认

为该系统的滞后时间约为 16 min. 经过 15 000 次学习之后,网络的训练误差为 0. 017 534,泛化误差为 0. 020 824. 图 5 给出了神经网络的输出同样本信号的对比曲线,从图中可以看出神经网络能表现出实际系统的幅值特性和延迟特性,能够对抽提塔底温度的变化进行实时预报. 本文所研究的方法为通过溶剂比调整抽提塔底温度提供了依据.

表 2 仿真条件(实例)

Tab. 2 Simulation conditions (example)

参数名称	参数值
网络结构	2-20-20-1
权值学习率	0. 001
延迟学习率	0. 001
节点函数	$f(x)=A\left(\frac{1-e^{-x}}{1+e^{-x}}\right)$
节点幅值 A	3
学习次数	15 000

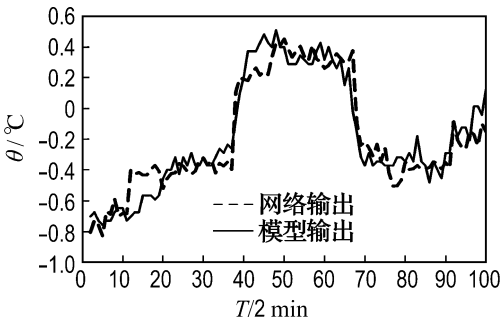


图 5 泛化信号同教师信号的比较(实例)

Fig. 5 Comparison between generalizing signal and teacher signal (example)

在仿真实验过程中,延迟时间的初值设定会对网络的学习效果有一定的影响. 因为在训练过程中,延迟时间参数学习率过大必然会导致延迟参数的调整频繁,会影响到系统的稳定性;而延迟时间参数过小会使网络学习效率降低,不能实现对被辨识对象延迟时间估计的作用. 一般可将输入层延迟时间参数的初始值设置为 1,输出层的延迟时间参数与被辨识对象的纯滞后时间相接近. 对于未知对象的辨识,可以粗略估计其延迟时间来进行网络参数的设置.

4 结 语

本文针对包含滞后的动态非线性系统辨识的要求和特点,提出了可以自适应选择网络延迟时间参数的动态 BP 算法. 该学习算法能够利用一个简单的网络结构在缺少先验知识的情况下,实现对非线性动态系统的建模. 仿真结果表明:本文

提出的动态BP网络能够很好地体现动态非线性系统的特性;能够根据网络的误差梯度对网络的延迟时间参数进行修正并且有利于网络精度的提高;最终辨识得到的网络延迟时间能够用于辨识对象的纯滞后时间.仿真实例证明本文所提出的动态BP网络改进算法为模型参数未知的非线性滞后系统辨识研究提供了新方法.

参考文献:

- [1] GOMEZ O, OROLOV Y, KOLMANOVSKY I V. On-line identification of SISO linear time-invariant delay system from output measurements [J]. *Automatica*, 2007, **43**(12):2060-2069
- [2] PEARSON R K. Outliers in process modeling and identification [J]. *IEEE Transactions on Control Systems Technology*, 2002, **10**(1):55-63
- [3] GÓMEZ J C, JUTAN A, BAEYENS E. Wiener model identification and predictive control of a pH neutralisation process [J]. *IEE Proceedings-Control Theory and Applications*, 2001, **151**(3):329-338
- [4] 王蓬,李少远.一类非线性系统的多模型预测控制[J]. *控制与决策*, 2007, **22**(10):1113-1118
- [5] 邢宗义,胡维礼,贾利民.基于T-S模型的模糊预测控制研究[J]. *控制与决策*, 2005, **20**(5):495-499
- [6] 张日东,王树青.基于支持向量机的一类非线性系统预测控制[J]. *控制与决策*, 2007, **22**(10):1103-1107
- [7] ABBAS H M, BAYOUMI M M. Volterra-system identification using adaptive real-coded genetic

- algorithm [J]. *IEEE Transactions on Systems, Man and Cybernetics — Part A:Systems and Humans*, 2006, **36**(4):671-684
- [8] BALESTRINO A, VERONA F B, LANDI A. On-line process estimation by ANNs and Smith controller design [J]. *IEE Proceedings-Control Theory and Applications*, 1998, **145**(2):231-235
- [9] 陆燕,杜继宏,李春文.延迟时间未知的延迟系统神经网络补偿控制[J]. *清华大学学报(自然科学版)*, 1998, **38**(9):67-69
- [10] REN X M, RAD A B. Identification of nonlinear systems with unknown time delay based on time-delay neural networks [J]. *IEEE Transactions on Neural Networks*, 2007, **18**(5):1536-1541
- [11] HAN M, HAN B, GUO W. Process control of pH neutralization based on adaptive algorithm of universal learning network [J]. *Journal of Process Control*, 2006, **16**(1):1-7
- [12] 邢建国,许沧粟,孙优贤.发动机怠速控制NARX模型及辨识[J]. *内燃机工程*, 2002, **23**(1):69-74
- [13] SIMON H. *Neural Networks: A Comprehensive Foundation* [M]. 2nd ed. Canada:Prentice Hall, 1999:162-167
- [14] MASTOROCOSTAS P A, THEOCHARIS J B. A recurrent fuzzy-neural model for dynamic system identification [J]. *IEEE Transactions on Systems, Man and Cybernetics — Part B: Cybernetics*, 2002, **32**(2):176-189

Nonlinear time delay systems identification based on dynamic BP algorithm

HAN Bing^{*1,2}, HAN Min¹

- (1. School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China;
2. Shanghai Ship and Shipping Research Institute, Shanghai 200135, China)

Abstract: The modeling of nonlinear dynamic system is an important problem in the domain of automatic control. For the problem, especially for the nonlinear time delay system, a novel dynamic BP (back propagation) neural network algorithm with adaptive time delay parameters is proposed. Based on the traditional multilayer perceptron neural network structure, the adaptive time delay parameters are employed in the first hidden layer and output layer. The gradient descent method helps to realize the parameter adjusting and time delay estimation. The simulation results show that the proposed method can not only identify the nonlinear system effectively, but also estimate time delay exactly.

Key words: nonlinear system; time delay; system identification; neural network