

连续搅拌反应釜(CSTR)的神经网络动态建模*

王永骥 徐桂英

(华中理工大学控制系 武汉 430074)

摘 要 基于广义性能指标,提出一种神经网络学习算法——广义递推预报误差学习算法(GRPE),该算法具有二阶收敛阶次。同时讨论了学习速率的选择问题。利用所提出方法对 CSTR 动态建模结果表明,基于 GRPE 训练的 DRNN 比基于 BP 训练的 MLP 模型精度高,收敛速度快。

关键词 广义性能指标,对角回归神经网络,递推预报误差学习算法,收敛性,连续反应搅拌釜

分类号 TP 183

Study on Neural Network Dynamic Modeling of Continuous Stirred Tank Reactor

Wang Yongji, Xu Guiying

(Huazhong University of Science and Technology)

Abstract Based on the generalized cost criterion, a new recursive prediction error learning algorithm (GRPE) for diagonal recurrent neural networks (DRNN) is presented, which has the second order convergence. A guideline for the choice of the optimal learning rate is derived from the convergence analysis. The application of this method to dynamic modeling of continuous stirred tank reactor (CSTR) shows that the GRPE trained DRNN has higher modeling precision and faster convergence than BP trained MLP.

Key words generalized cost criterion, diagonal recurrent neural networks, recursive prediction error learning algorithm, convergence, continuous stirred tank reactor (CSTR)

1 引 言

许多化工过程都是多变量非线性系统,连续搅拌反应釜(CSTR)也是如此。为了提高 CSTR 的控制性能,必须采用先进的控制技术。首要一步是建立过程的动态数学模型,为此,本文研究采用神经网络进行 CSTR 的动态建模问题。

利用神经网络建模的主要问题是选择神经网络的类型和结构,其次是训练^[1]。由于动态神经网络比静态神经网络在系统控制方面具有更大的灵活性,且结构简单,本文采用对角回归神经网络(DRNN)^[2]进行建模。

分析表明,神经网络训练的收敛速度不仅取决

于所采用的优化算法,而且与待优化的性能指标有关。基于这一思想,文献[3]介绍了一种广义性能指标。该指标是 l_1 范数性能指标(绝对值误差和准则)和 l_2 范数性能指标(二次性能指标)的线性组合,基于该性能指标训练神经网络,可加快收敛速度,具有较强的鲁棒性。

基于广义性能指标,本文提出一种用于 DRNN 的广义递推预报误差学习算法(GRPE),并讨论了算法的收敛性问题。结果表明,基于 GRPE 训练的 DRNN 用于建立 CSTR 的动态模型,与基于 BP 训练的 MLP 相比,DRNN 具有更高的建模精度,且需要的计算时间较少。

* 国家自然科学基金(69974017)和工业控制技术国家重点实验室开放课题基金项目(K97M01)

2 对角回归神经网络

对角回归神经网络的示意图如图 1 所示。

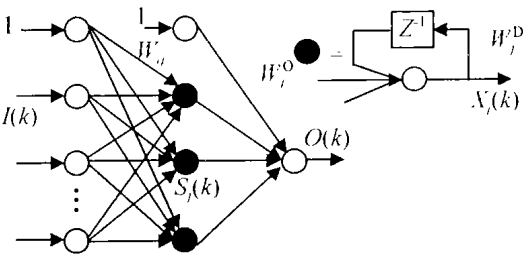


图 1 对角回归神经网络结构示意图

图中,黑圆表示回归神经元,白圆表示线性元。对每一离散时刻 k , $I_i(k)$ 是第 i 个输入, $S_j(k)$ 是第 j 个回归元的输入总和, $X_j(k)$ 是第 j 个神经元的输出, $O(k)$ 是回归神经网络的输出, W , W^O 和 W^D 分别表示输入、输出和对角反馈加权向量, R^{ni} , R^{nh} , R^{no} 分别为它们的维数,ni,nh,no分别是输入层、隐含层和输出层神经元的个数。

DRNN 的输入输出关系式为

$$\begin{cases} O(k) = \sum_j W_j^O X_j(k) \\ X_j(k) = f(S_j(k)) \\ S_j(k) = W_j^D X_j(k-1) + \sum_i W_{ji} I_i(k) \end{cases} \tag{1}$$

$$f(x) = \frac{1 - \exp(x)}{1 + \exp(x)} \tag{2}$$

3 基于广义性能指标的递推预报误差学习算法

广义性能指标的评价函数 $G(\lambda)$ 定义为^[3]

$$G(\lambda) = \lambda E + (1 - \lambda) E' = \lambda \sum_{k=1}^N \Phi_2(E(k)) + (1 - \lambda) \sum_{k=1}^N \Phi(E(k)) \tag{3}$$

其中, E 为预报误差向量, $\Phi_2(E) = \frac{1}{2}E^T E$ 为误差平方和, $\Phi(\cdot)$ 为正定的、凸的连续可微函数, $\lambda \in [0, 1]$ 为加权系数,在学习过程中是变化的, N 是样本数。

评价函数保证在优化过程初期具有较快的收敛性, $G(0) = E'$ 的选择对优化过程后期的收敛速度影响很大。选择

$$\Phi(x) = \beta \ln[\cosh(x/\beta)] \tag{4}$$

其中 β 是控制参数。分析表明^[3], β 较大时, $G(0)$ 近似于二次性能指标准则(l_2 范数准则); β 较小时,

$G(0)$ 近似于绝对值准则(l_1 范数准则)。众所周知,二次性能指标仅适合于正态高斯噪声的情况。如果系统噪声为非高斯噪声,特别是非正规的尖峰噪声时,基于二次性能指标的学习结果往往不够理想。但是,函数(3)可适用于非高斯噪声的情况,即它具有一定的鲁棒性。

递推预报误差学习算法(RPE)是一种参数估计算法,它极小化(5)式定义的性能指标函数

$$\begin{cases} J(\theta) = \frac{1}{2N} \sum_{j=1}^N E^T(j, \theta) E(j, \theta) \\ \theta(k) = [W_j^O(k) \quad W_j^D(k) \quad W_{ij}(k)]^T \end{cases} \tag{5}$$

其中, $E(j, \theta)$ 是具有 N 个输出的系统预报误差向量。定义 $J(k)$ 为神经网络输出 $O(k)$ 对 $\theta(k)$ 的偏导数,则采用广义性能指标时,本文提出的GRPE算法如下

$$\begin{cases} e(k) = y(k) - O(k) \\ P(k) = \frac{1}{\alpha(k)} \{ P(k) - c(k) P(k) J(k) [\alpha(k) + c(k) J^T(k) P(k) J(k)]^{-1} J^T(k) P(k) \} \\ c(k) = \{ \lambda(k) + \frac{1}{\beta} (1 - \lambda(k)) [1 - \tanh(e(k)/\beta)] \} \\ g(k) = \{ \lambda(k) e(k) + (1 - \lambda(k) \tanh[e(k)/\beta] J(k) \} \\ \lambda(k) = \exp(-1/E(k)) \\ E(k) = \frac{1}{2} \sum_{i=1}^k e^2(i) \\ \theta(k) = \theta(k-1) + \eta(k) P(k) g(k) \\ J(k) = [X_j(k) \quad W_j^O L_j(k) \quad W_j^O Q_{ij}(k)] \end{cases} \tag{6}$$

用于 DRNN 的 GRPE 在线更新算法为:

Step1: 随机初始化加权矩阵, $P(0) = 10^4 \mathcal{G}\mathcal{G}$ 为单位矩阵,选择 $\alpha(0)$, $\lambda(0)$, $\eta(0)$ 的适当值;

Step2: 根据 DRNN 的输入,利用(1),(2)式计算 $S_j(k)$, $X_j(k)$ 和 $O(k)$;

Step3: 根据(6)式构造定义的 J 向量,计算预报误差向量 $e(k)$,矩阵 $P(k)$ 及参数向量 $\theta(k)$;转Step2,直到满足收敛。

准则参数更新规则要求适当选择学习速率 $\eta(k)$ 。GRPE 的收敛性类似于 QRPE^[4],基于 Lyapunov 稳定性分析,这里给出一种选择 η 的方法。

$$0 < \eta < 2/g_{\max} \tag{7}$$

其中

$g_{\max} = \max |g^m(k)| = \max |J^T(k)Q(k)J(k)|$
 $Q(k)$ 为与训练过程有关的正定矩阵^[4]。

4 连续搅拌反应釜的动态建模

连续搅拌反应釜(CSTR) 是一多变量非线性系统,为了提高其控制性能,必须采用先进的控制技术。首先利用神经网络对 CSTR 进行动态建模,同时将 GPRE 训练的 DRNN 和 BP 以及 QRPE 训练的 DRNN 进行比较。醋酸在 CSTR 中与腐蚀性的碱中和,由于该过程是强非线性的,用于建模的输入/ 输出数据来源于闭环 PI 控制^[5],输出流中 pH 值和腐蚀性碱泵的频率关系为

$$\text{pH}(k) = f(\text{pH}(k - 1), \dots, \text{pH}(k - n), f(k - 1), \dots, f(k - m)) \tag{8}$$

其中 n 和 m 分别为系统的输出和输入的动态阶次。DRNN 结构选为 3 - 6 - 1, 输入层考虑偏置,此时输入为

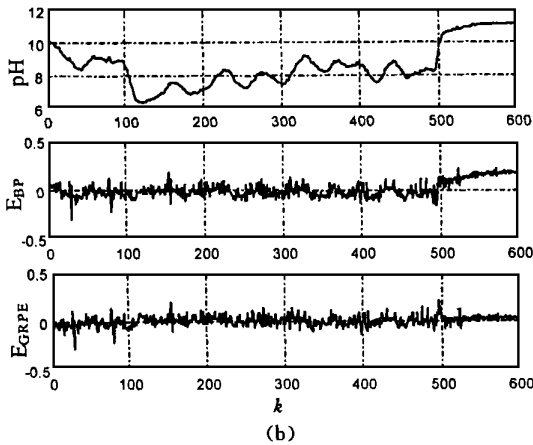
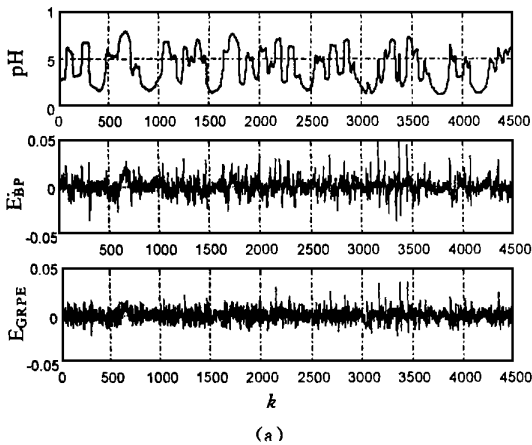


图 2 检验结果

(a) 5 步预报精度(训练数据)

(b) 1 步预报精度(测试数据)

$I(k) = [\text{pH}(k), f(k), 1]^T \tag{9}$

加权系数的初值在 - 1 和 1 之间随机选择。 η 根据式 (7) 在线更新。

基于 BP 算法训练 MLP 网络时,网络为一个包含 2 个隐含层的 4 层网络,其结构为 11 - 8 - 8 - 1^[5]。BP 算法的更新规则为

$$\Delta w_{ij}(k) = \alpha_j \frac{\partial E}{\partial w_{ij}} + \beta \Delta w_{ij}(k - 1) \tag{10}$$

其中, α_j 根据 S A B 法则^[5] 更新, $\alpha_j(0) = 0.07, \beta = 0.02$ 。

为了检验训练好的神经网络的正确性,将测试数据用于校验网络,检验 1 步和 5 步预报误差。图 2 给出了检验结果。

表 1 给出了 MLP 和 GRPE 用于 CSTR 的建模结果。

表 1 BP - MLP 和 GRPE - DRNN 建模精度的比较

		BP - MLP	GRPE - DRNN
1 步预报误差	训练	5.867 6E - 5	3.085 3E - 5
	测试	0.006 6	0.002 9
5 步预报误差	训练	9.283 0E - 4	3.053 7E - 4
	测试	0.048 1	0.031 1

5 结 语

误差的评价准则是影响神经网络学习收敛速度的一个重要因素。广义性能指标函数综合了 l_2 和 l_1 范数性能指标,保证了训练初期和后期都具有较快的收敛速度。基于广义性能指标函数,本文提出一种广义预报误差训练算法,将该算法用于对角回归神经网络,具有二阶收敛速度。用本文算法对 CSTR 动态建模,结果表明效果良好,训练速度快并且精度高。

致谢 本文的部分工作是第一作者在德国多特蒙德大学高访时完成的。为此,作者谨对 Engel 教授和 Fernholz 博士的支持表示感谢。

参 考 文 献

1 王永骥,涂建. 神经网络控制. 北京: 机械工业出版社, 1998
2 Ku C C, Lee K Y. Diagonal recurrent neural networks for dynamic systems control. IEEE Trans on Neural Networks, 1995, 6(1): 144- 156
3 Karayiannis N B, Venetsanopoulos A N. Artificial neural networks, learning algorithms, performance evaluation and applications. Kluwer Academic Publishers, 1993

近一般PID控制器输出,增加了辨识的稳定性。在实际过程中发现,辨识器一般在辨识几步后即达到稳定,这种情况的控制效果比较理想,接近于仿真曲线。

参 考 文 献

1 David J Kruglinski 著. 王国印译. Visual C+ + 技术内幕. 第二版. 北京: 清华大学出版社, 1996

2 M F Arnett, M Coulombe. TCP/IP 实用技术指南. 北京: 清华大学出版社, 1997

3 Martin Hall. Windows sockets——An open interface for network programming under Microsoft windows. USA: Winsock Group, 1993

4 Martin Hall *et al.*. Windows sockets 2 application programming interface——An interface for transparent network programming under Microsoft windows (Revision 2.2.1), USA: Winsock Group, 1997

5 舒迪前, 饶立昌, 柴天佑. 自适应控制. 沈阳: 东北大学出版社, 1993

作 者 简 介

张国范 女, 1944年生。东北大学副教授。主要从事计算机控制、管理与控制相结合分级系统的教学与科研工作。

董视秋 女, 1955年生。沈阳工业学院副教授。主要从事计算机网络及智能控制的教学与科研工作。

(上接第 377 页)

4 Yongji Wang, Gregor Fernholz, Sebastian Engell. A second order recursive prediction error algorithm for diagonal recurrent neural networks. In: Proc of IEEE Conf on Control Application. Trierot, 1998. 172– 176

5 Draeger A. Neuronale Netze zur Regelung verfahrenstechnischer Anlagen. Doctor Dissertation, Universitat Dortmund, Shaker Verlag, 1997

作 者 简 介

王永骥 男, 1955年生。1990年于华中理工大学获博士学位, 1997– 1998年在德国多特蒙德大学化工系高访, 现为华中理工大学控制系教授, 美国IEEE会员。主要研究方向为神经网络控制, 预测控制, 工业过程计算机控制应用。

徐桂英 女, 1945年生。1970年毕业于华中理工大学。现为该校控制系副教授, 中国自动化学会理事。主要研究方向为神经网络控制, 工业过程计算机控制应用。