

循环神经网络建模在非线形预测控制中的应用*

古 勇 苏宏业 褚 健

(浙江大学工业控制技术国家重点实验室、先进控制研究所 杭州 310027)

摘 要 基于动态 Levenberg- Marquardt (LM) 算法, 提出两步 LM 方法建立非线性过程的循环神经网络模型。该模型能以足够的精度并行于过程运行, 并能从过程的输入信息模拟过程未来的响应。研究了基于该模型的扩展 DM C 预测控制策略, 仿真结果表明该控制器的性能得到了很大提高。

关键词 循环神经网络, LM 算法, DM C 预测控制

分类号 TP 264

Recurrent Neural Network Modeling and Its Application in Nonlinear Predictive Control

Gu Yong, Su Hongye, Chu Jian

(Zhejiang University)

Abstract A two step Levenberg- Marquardt (LM) algorithm for designing a recurrent neural network process model was proposed based on dynamic LM algorithm. The obtained process model is sufficiently accurate to be used independently from the process, emulating the process response from only process input information. Implementation of an extend DM C predictive control strategy based on the identified recurrent neural network model was studied, and the simulation result shows the improvements in control performance.

Key words recurrent neural network, LM algorithm, DM C predictive control

1 引 言

预测控制已成为过程控制的主要途径之一, 预测控制器的设计依赖于描述过程动态特性的过程模型。由于已证明多层前向神经网络能以任意精度逼近非线性输入输出映射^[1], 因此神经网络被广泛用于动态非线性过程的建模。过程动态可通过将输入神经元的输入取为时延后的过程输入输出而引入神经网络模型。若其中的过程输出由相应时刻的网络模型输出取代, 则这样的神经网络为循环神经网络模型。

Werbos^[2] 提出用 Backpropagation Through Time (BPTT) 算法来训练循环神经网络, 这种算法能准确有效地计算循环神经网络训练中的导数。BPTT 算法源自非线性最小二乘优化算法中的最速下降法, 它收敛很慢, 而且有可能很快收敛到某一局

部极小点, 过程模型的预测精度也很差。LM 算法为改进的 Gauss- Newton 算法, 比最速下降法收敛性能好且收敛速度快^[3]。

本文将 LM 算法推广到动态情形, 提出两步 LM 算法, 该算法能以较快的速度得到具有更高精度的循环神经网络过程模型。采用该模型作为扩展 DM C 控制策略^[4]的预测模型, 并应用于典型连续搅拌反应釜对象的控制。仿真结果表明, 动态 LM 算法可提高扩展 DM C 控制器的控制性能。

2 基于动态 LM 算法的循环神经网络两步 LM 算法

假设非线性动态系统可用非线性离散差分方程来描述, 即

$$\begin{aligned} y^p(t+1) = & \\ f[y^p(t), \dots, y^p(t-n_y+1), & \\ u(t), \dots, u(t-n_u+1)] & \end{aligned} \quad (1)$$

* 1998- 10- 09 收稿, 1999- 01- 04 修回

其中, $y^p(t) \in R^n, u(t) \in R^m$ 分别为系统的输出和输入, n_y 和 n_u 为系统的阶次。基于神经网络建模, 网络模型可表示为两种形式, 即

$$\begin{aligned} \hat{y}^m(t+1) = & f[y^p(t), \dots, y^p(t-n_y+1), \\ & u(t), \dots, u(t-n_u+1), w] + e(t) \quad (2) \end{aligned}$$

$$\begin{aligned} \hat{y}^m(t+1) = & f[y^m(t), \dots, y^m(t-n_y+1), \\ & u(t), \dots, u(t-n_u+1), w] + e(t) \quad (3) \end{aligned}$$

其中, y^m 为神经网络模型的输出, f 为神经网络逼近的非线性输入输出映射。注意到(3)式的自变量中包含网络本身过去时刻的输出值, 此即循环神经网络模型结构。

一般神经网络权值训练的LM算法形式参见文献[3]。该算法的关键是如何求残差 $r(t, w)$ 对网络权值的导数。对于(2)式的模型结构, 此导数的计算十分简单, 即

$$\frac{dr(t, w)}{dw_j} = \frac{\partial y^m(t, w)}{\partial w_j} = \frac{\partial \hat{y}^m(t, w)}{\partial w_j} \quad (4)$$

本文针对(3)式的循环神经网络模型结构, 采用链式规则计算其导数, 即

$$\begin{aligned} \frac{dr(t, w)}{dw_j} = \frac{dy^m(t)}{dw_j} = & \frac{\partial y^m(t)}{\partial w_j} + \sum_{\tau=1}^{n_y} \frac{\partial y^m(t)}{\partial y^m(t-\tau)} \frac{dy^m(t-\tau)}{dw_j} \quad (5) \end{aligned}$$

当循环神经网络模型以过程 k 时刻状态为基础, 做 s 步长期预测时, 长期预测残差 $r(t+s, w)$ 对网络权值的导数可通过以下步骤计算:

Step 1: 当 $t = k$ 时, 置 $y^m(t) = y^p(t)$, 则 $dy^m(t)/dw_j = 0$;

Step 2: 置 $t = k + 1$;

Step 3: 根据神经网络的映射表达式计算 $\frac{\partial y^m(t)}{\partial w_j}$ 和 $\frac{\partial y^m(t)}{\partial y^m(t-\tau)}$;

Step 4: 根据(5)式计算 $\frac{dr(t, w)}{dw_j} = \frac{dy^m(t)}{dw_j}$;

Step 5: 当 $t = k + s$ 时停止; 否则令 $t = t + 1$, 转 Step 3。

我们称针对(2)式模型结构的LM算法为静态LM算法; 称针对(3)式循环神经网络模型结构的LM算法为动态LM算法。由此提出用两步LM算法来训练过程的神经网络模型。第一步神经网络模型采用(2)式的模型结构, 当静态LM算法达到局部极小点后, 神经网络转成循环神经网络模型结构, 再采用动态

LM算法进行训练。仿真例子表明, 动态LM算法可在静态LM算法的基础上进一步减小测试样本集上的长期预测误差。

3 循环神经网络两步LM法建模

考虑连续搅拌釜反应器中发生的一阶放热不可逆反应^[5], 其动态方程的无量纲形式为

$$\begin{aligned} dx_1/dt = & -x_1 + Da(1-x_1) \times \\ & \exp\left(\frac{x_2}{1+x_2/\gamma}\right) = f_1(x_1, x_2) \quad (6) \\ dx_2/dt = & -x_2 + BDa(1-x_1) \times \\ & \exp\left(\frac{x_2}{1+x_2/\gamma}\right) - \beta(x_2 - u_c) = \\ & f_2(x_1, x_2, u_c) \quad (7) \end{aligned}$$

其中 x_1, x_2 和 u_c 为无量纲的产物浓度, 反应温度和冷夹套温度; 控制变量为产物浓度 x_1 , 操作变量为冷夹套温度 u_c ; Da, γ, B 和 β 为已知物理参数。

在建立过程的循环神经网络模型时, 首先采用多值伪随机信号作为过程的激励输入, 采样间隔取 0.4 倍无量纲单位时间。从 CSTR 过程的响应看, 过程的二阶动态占主导地位, 而且没有明显的时滞, 因此模型的阶数取 2, 即 $n_y = 2, n_u = 2$ 。从网络泛化能力考虑, 能逼近输入输出数据样本集的最小结构的网络一般都具有最好的泛化能力。通过仿真发现, 4 个隐层节点数是网络逼近精度和泛化能力的最佳折衷。因此得到的神经网络的拓扑结构为: 4 个输入节点, 4 个隐层节点和 1 个输出节点。

为模拟真实过程数据测量噪声的影响, 假设在过程输出附加均匀分布的白噪声, 噪声方差为输出幅值的 5%。在神经网络的训练过程中, 为避免过训练, 需不断监控测试样本集上每预测 8 步的长期预测误差平方和。当此误差不再下降并开始上升时, 训练从静态LM算法转为动态LM算法或终止。

经分析, 静态LM算法在第 40 回合时测试误差不再下降并开始上升, 因此在第 40 回合得到神经网络模型 A, 然后网络改为循环神经网络结构, 采用动态LM算法。长期预测步长取 20 步, 即每过 20 个采样时刻, 重置网络模型输入向量中的输出项为过程各时刻的输出值。训练至 85 回合时终止, 得到神经网络模型 B。

可以看出, 动态LM算法可进一步减小测试样本上的长期预测误差, 这是因为动态LM算法从本质上考虑了循环神经网络的循环反馈结构。为深入比较神经网络模型 A 和 B 的动态特性, 在图 1 中给

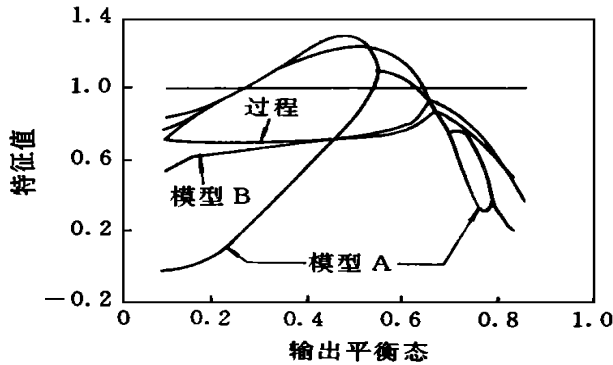


图1 过程与模型局部线性化的特征根

出了过程和模型在稳态点上局部线性化的两特征根相对稳态输出曲线, 两曲线重合表明它们是一对共轭复根的实部。显然, 模型B比模型A更接近过程的特征根曲线, 说明模型B在动态本质上更逼近过程。

4 基于循环神经网络的扩展DMC预测控制算法

基于上节得到的神经网络过程模型A和B, 结合 Peterson^[4]的扩展DMC预测控制策略, 控制CSTR过程从低平稳平衡态过渡到中间非平稳平衡态。具体的扩展DMC算法参见文献[4]。

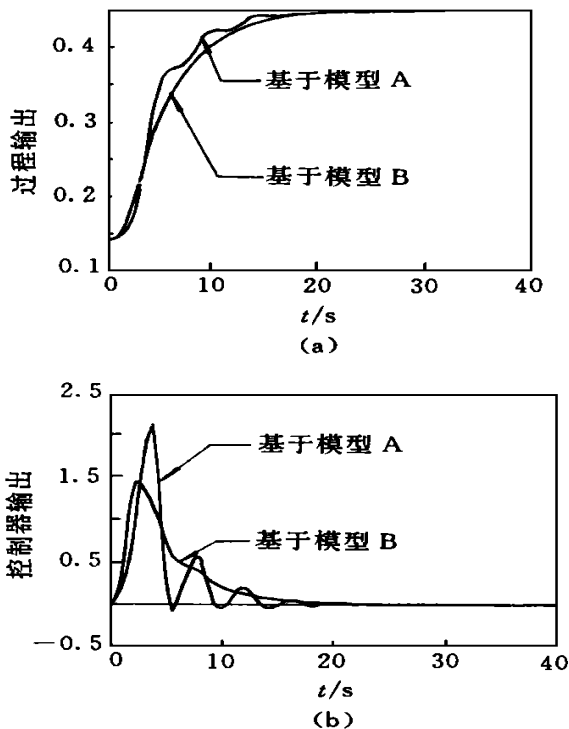


图2 扩展DMC预测控制下的过程输出响应和控制器输出
(a) 过程输出 (b) 控制器输出

基于模型的预测控制器性能的好坏与模型的精度密切相关。图2显示了基于模型A和B的过程输出响应及控制器输出, 其中控制步长取 $M=2$, 预测步长取 $P=5$ 。可以看出, 基于循环神经网络B的控制效果比基于模型A的效果有很大改善。由于模型B长期预测的准确性, 反馈校正所起作用不大, 控制器输出十分平稳, 控制输出的最大幅值也得以减小, 过程的输出响应变得十分平稳。

5 结论

本文探讨循环神经网络在非线形动态过程建模和控制中的应用, 提出两步LM算法训练循环神经网络模型。在非线形CSTR过程的建模应用中, 第二步动态LM算法能进一步降低模型的长期预测误差, 提高网络模型的泛化推广能力。从网络的局部线性化模型的特征根可以看出, 采用动态LM算法能更准确地逼近过程的动态特性。本文结合扩展DMC预测控制策略来控制开环不稳定CSTR对象。仿真结果表明, 基于动态LM算法的循环神经网络模型可以大大提高控制器的控制平稳性能。

参考文献

- 1 Hornik K, M Stinchcombe, H White. Multilayer feed-forward networks are universal approximator. *Neural Networks*, 1989, 2: 359- 366
- 2 Werbos P J. Backpropagation through time: What it does and how to do it? *Proc of the IEE*, 1990, 78(10): 1550- 1560
- 3 McKeown J J, Stella F, Hall G. Some numerical aspects of the training problem for feed-forward neural nets. *Neural Networks*, 1997, 10(8): 1455- 1463
- 4 Peterson T, Hernandez E, Arkun Y. A nonlinear DMC algorithm and its application to a semibatch polymerization reactor. *Chem Eng Sci*, 1992, 47(4): 737- 753
- 5 Uppal A, Ray W H, Poore A B. On the dynamic behavior of continuous stirred tank reactors. *Chem Eng Sci*, 1974, 29(5): 967- 985

作者简介

古勇 1972年生。1996年于浙江大学电机系获硕士学位, 现为浙江大学工业控制技术研究所博士研究生。主要研究方向为非线性系统辨识技术, 神经网络在工业控制中的应用, 非线性系统预测控制。

苏宏业 见本刊1996年第11卷第5期第604页。

褚健 见本刊1996年第11卷第5期第604页。