

RNN 神经网络的应用研究^{*}

朱群雄 孙 锋

(北京化工大学计算机系, 北京, 100029)

摘 要 对动态回归神经网络模型结构与算法进行了分析, 采用多层反馈 RNN 网络, 以典型的非线性化工过程 CSTR 为应用对象, 比较了采用前馈 BP 网络和 Elman 的 RNN 网络进行模型化与模拟, 最后用一个时变过程和苯酐工业生产过程模拟验证。结果表明, 动态回归神经网络具有较好的收敛性和稳定性, 可用于复杂动态过程的工业应用。

关键词 回归神经网络; Elman 网络; BP 网络; 反应过程

分类号 TP18

神经网络理论的应用现在已经渗透到各个领域, 特别是前馈网络 BP 算法在过程模拟、控制、优化操作等方面取得了较显著的成绩。前馈神经网络从学习观点看, 是一种强有力的学习系统, 其结构简单; 从系统观点看, 它是一种静态非线性映射, 通过简单非线性处理单元的复合映射可获得复杂的非线性处理能力。而从计算观点看, 前馈神经网络不是一种强有力的动力系统, 缺乏丰富的动力学行为。为了使网络成为一个非线性动力系统, 在 80 年代末, 国际上一些神经网络专家如 Jordan, Pineda, Williams, Elman 等提出了一种新神经网络结构模型, 即回归神经网络 (Recurrent Neural Network, RNN)。这种网络的本质特征是在处理单元之间既有内部的反馈连接又有前馈连接。从系统观点看, 它是一个反馈动力系统, 在计算过程中体现过程动态特性, 比前馈神经网络具有更强的动态行为和计算能力。回归神经网络现已成为国际上神经网络专家研究的重要对象之一, 本文针对化工连续动态过程特点, 分析 RNN 网络与算法并进行应用研究。

1 回归神经网络结构分析

从原理分析可知, 回归神经网络的内部反馈连接可以为网络保留隐层节点的状态和提供记忆方式, 网络的输出不再仅仅简单地取决于当前的输入, 而且和以前的网络内部状态有关, 体现出其较好的动态特性。目前, 国际上现有 10 种左右的回归神经网络模型, 应用最广泛的是 Elman 网络, 它与标准的前馈网络相似, 同样有输入、输出、隐含单元, 但前者还包括一组反馈单元, 它们用来保存隐含层单元的上个时刻的输出状态, 以表征对象的内部特征。这些内部状态与外部输入信号一起作为当前时刻隐层节点的输入。Karjala^[1] 等人在 1992 年已成功地将 Elman 网络运用在有高斯噪音条件下过程测量的数据修正

收稿日期: 1997-07-01

第一作者: 男, 37 岁, 博士, 副教授

^{*} 国家自然科学基金资助项目 (No. 29676002)

©1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

上, 证明 Elman 网络适合于在有噪音存在的情况下学习过程特性, 并且能显著地降低噪音对学习对象的影响。Yong You 和 Michael Nikolaou^[2] 已将 Elman 网络应用在非线性静态和动态过程模型的建立上。结论表明, 利用回归神经网络建立对象模型, 可以有效地减小输入空间, 特别是对于多变量的对象而言, 回归神经网络的结构比前向网络更接近实际模型。但是, 对于一些具有严重动态时变性、非线性、高阶次性等特点, 并且其过程中间状态变量较多的对象, 用 Elman 网络就很难表征其过程。例如, 在化工过程中, 由于处理组分较多, 各种影响因素又相互关联, 存在极强的非线性关系。目前对于这些过程的机理虽不乏研究, 但面临实际生产中众多问题, 如: (1) 反应机理的复杂性和大量物性数据的缺乏; (2) 工艺流程的复杂性及前后流程具有严重的反馈和关联耦合; (3) 众多随机因素。为此本文采用多层反馈 RNN 网络, 结构如下图所示:

增加多层反馈子层可以保留若干个时刻的网络内部状态, 以便更好地表征过程对象的内部特征。反馈子层 i 保留第 $(t-i)$ 时刻的隐含层节点的输出, 所以在网络中可以保存 n 个时刻的隐含层节点的输出。在第 t 个时刻, 第 i 反馈子层的内容将送入第 $(i+1)$ 反馈子层 ($i=1, \dots, n$), 而反馈子层 1 的内容将被隐含层单元

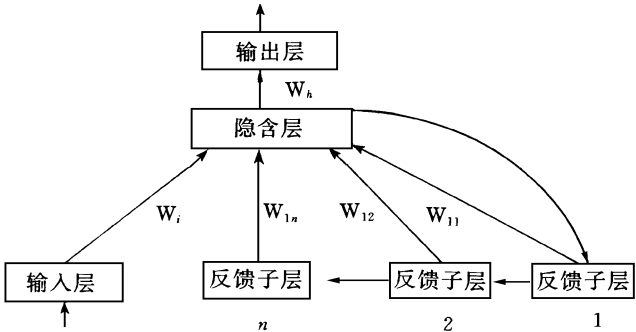


图 1 多层反馈 RNN 网络

的输出而取代。各反馈子层可进行短时的记忆, 以保存一些必要的网络内部信息以表征过程的一些空间状态变量。与原始的 Elman 网络相比较, 多层反馈 RNN 网络保存了更多的过程状态信息, 更加接近实际的模型, 表现出更强的动态特性, 可以对更复杂的对象进行模拟和控制。

本文采用基于梯度下降的 BP 算法来训练回归神经网络, 节点函数为 Sigmoid 函数, 设 $\overline{y(t)}$ 为网络的实际输出, $y(t)$ 为理想值, 则权值的调整公式为:

$$\Delta w_{ij} = - \eta \frac{\partial E}{\partial W_{ij}} \tag{1}$$

当 j 是输出层节点时:

$$\frac{\partial E}{\partial W_{jk}} = - (y(t) - \overline{y(t)}) f(U_k) (1 - f(U_k)) V_j \tag{2}$$

当 j 是隐含层节点时:

$$\frac{\partial E}{\partial W_{ij}} = - (y(t) - \overline{y(t)}) f(U_k) (1 - f(U_k)) W_{jk} f(U_j) (1 - f(U_j)) V_i \tag{3}$$

同理可对反馈层节点的权值进行调整, 即:

$$\frac{\partial E}{\partial W_{lj}} = - (y(t) - \overline{y(t)}) f(U_k) (1 - f(U_k)) W_{jk} f(U_j) (1 - f(U_j)) V_l \tag{4}$$

式中 U_k 和 V_k 分别为第 k 个节点的输入和输出。

2 连续搅拌釜反应过程应用研究

连续搅拌釜反应器 CSTR 的过程动态关系^[3]可用下列方程表示:

型

$$\frac{dC_A}{dt} = \frac{Q}{V}(C_{A,f} - C_A) - k_0 C_A \exp\left[-\frac{E}{kT}\right]$$
$$\frac{dT}{dt} = \frac{Q}{V}(T_f - T) - \frac{(-\Delta H)k_0 C_A}{\rho C_p} \exp\left[-\frac{E}{kT}\right]$$

近实

(5)

馈子

保

$$\frac{P_c C_{pc}}{\rho C_p V} Q_c \left[1 - \exp\left[-\frac{h_A}{Q_c \rho_c C_{p,c}}\right] \right] (T_{c,f} - T)$$

(6)

上式中的参数值如表 1 所示。

表 1 CSTR 方程参数表

$Q/L^{\circ}\text{min}^{-1}$	$C_{A,f}/\text{mol}^{\circ}\text{L}^{-1}$	$T_f, T_{c,f}/\text{K}$	V/L	$h_A/\text{J}^{\circ}\text{min}^{-1}\cdot\text{K}^{-1}$
100	1	350	100	2.93×10^6
k_0/min^{-1}	$E^{\circ}\text{kJ}^{\circ}\text{K}^{-1}$	$-\Delta H/\text{J}^{\circ}\text{mol}^{-1}$	$\rho/\text{g}^{\circ}\text{L}^{-1}$	$\rho_c/\text{g}^{\circ}\text{L}^{-1}$
7.2×10^6	9.95×10^3	8.36×10^5	1 000	1 000
$C_p/\text{J}^{\circ}\text{g}^{-1}\cdot\text{K}^{-1}$	$C_{p,c}/\text{J}^{\circ}\text{g}^{-1}\cdot\text{K}^{-1}$	$Q_c/L^{\circ}\text{min}^{-1}$	T/K	$C_A/\text{mol}^{\circ}\text{L}^{-1}$
4.18	4.18	103.41	440.2	8.36×10^{-2}

冷却水流量 Q_c 影响反应温度 T ，而温度 T 影响组分 A 的浓度 C_A 。现在要建立浓度 C_A 与冷却水流量 Q_c 之间的关系。表 2 给出了用 BP 网络（1-4-1）、Elman 网络（1-3-1，一个反馈层）和本文所用的多层反馈 RNN 网络（1-3-1，二个反馈子层）训练模拟结果。由此可以看出，采用多层反馈回归神经网络具有较好的收敛精度和适应性。根据多次实践，我们认为反馈子层以二层为宜。

表 2 神经网络训练模拟结果比较

项 目	基于训练样本		基于未训练样本		同一计算机稳定收敛训练时间/s
	平均绝对误差	平均相对误差/%	平均绝对误差	平均相对误差/%	
BP 网络	0.000 104	0.118 988	0.000 102	0.117 323	41
Elman 网络	0.000 105	0.120 669	0.000 104	0.119 468	60
多层反馈 RNN 网络	$7.475\,499\times 10^{-5}$	0.085 886	$7.364\,61\times 10^{-5}$	0.084 360	65

3 时变过程的神经网络模型

由澳大利亚纽卡斯尔大学的 Graebe 博士设计的控制对象是一个单入单出的时变系统^[4]，闭环控制系统的设定值规定是一个周期为 20 s 的幅值为±1 的方波。要求被控输出在所有时间内都必须在±1.5 范围内，在大多数时间内，被控输出值与设定值的偏差绝对值最好小于 0.2，偶然的绝对偏差大于 0.2 是允许的，调节速度越快越好。本文建立了被控参数和控制信号为输入、被控输出值为输出的带 2 个反馈子层的 2-6-1 结构的控制过程回归神经网络模型。图 2 为基于训练样本的模拟结果，其平均绝对误差为 0.022 2，图 3

为基于未训练的测试样本的模拟结果, 其平均绝对误差为 0. 024 1。

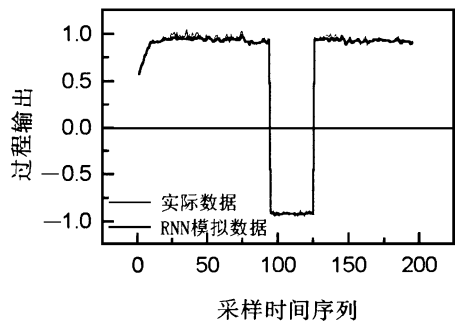


图 2 基于训练样本的模拟结果

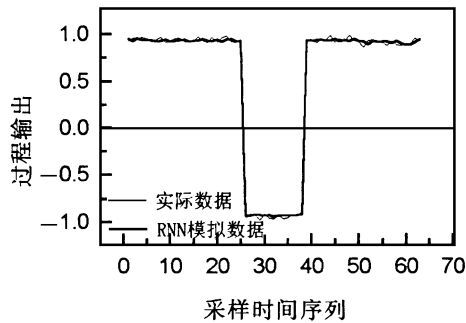


图 3 基于未训练的测试样本的模拟结果

4 苯酐生产反应过程应用实践

萘催化氧化流化床反应器是苯酐生产系统的关键设备, 其反应机理和流化床模型较为复杂, 副产品较多, 有萘醌、顺丁烯二酸酐、CO₂、CO 和 H₂O 等。影响苯酐选择性的因素较多, 可操作参数有反应温度、空萘比、送风量等^[5]。现在从某生产厂采集一批样本数据, 用多层反馈回归神经网络建立其模型, 设网络结构为 3 入 4 中 1 出, 2 个反馈子层。图 4 为基于训练样本的模拟结果, 其平均绝对误差为 1. 090 7, 平均相对误差为 1. 396 1%, 图 5 为基于未训练的测试样本的模拟结果, 其平均绝对误差为 1. 118 1, 平均相对误差为 1. 426 3%。

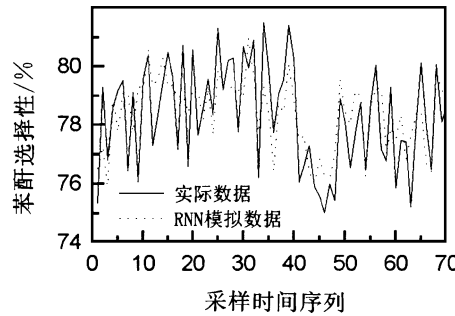


图 4 基于训练样本的模拟结果

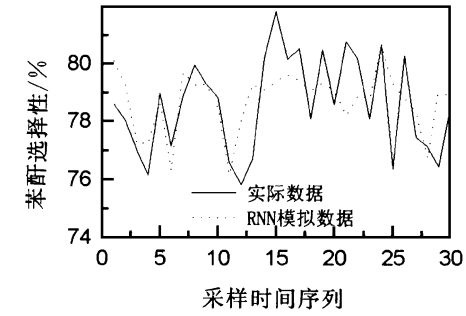


图 5 基于未训练的测试样本的模拟结果

5 结 论

有双反馈子层的回归神经网络能更好地反映对象过程中的状态空间, 体现出较强的动态特性, 模型结构更趋合理, 网络收敛性和预测精度高。实践证明, RNN 网络为复杂动态过程模拟和控制提供了一种新的有效的建模方法。

参 考 文 献

- 1 Karjala T W, Himmelblan D M, Mäikkiläinen R. Data rectification using recurrent (Elman) neural networks. Proc IEEE-IJCNN, Baltimore, 1992
- 2 You Y, Nikolaou M. Dynamic process modeling with recurrent neural networks. AIChE J. 1993, 39: 1654
- 3 Henson M A, Seborg D E. Input-output linearization of general nonlinear process. AIChE J. 1990, 36: 1753~1757
- 4 Graebe S F. Robust and adaptive control of an unknown plant: A benchmark of new format. Automatica, 1994, 30 (4): 567~575
- 5 朱群雄, 麻德贤, 阚丹峰. 蔡催化氧化流化床反应器的模型化和优化计算. 计算机与应用化学, 1990, 7 (4): 254~258

Study on application of recurrent neural network

Zhu Qunxiong Sun Feng

(Department of Computer, Beijing University of Chemical Technology, Beijing, 100029)

Abstract Recurrent neural networks (RNNs) have been widely used in the dynamic process modeling, simulation and control. In this paper the RNN structure and algorithm are analyzed. The modified RNN network used is compared with the BP network as well as Elman network by application of CSTR modeling and simulation. The industrial reaction process of phthalic anhydride production and a benchmark of new format are demonstrated using the RNN with two context sublayers. These results exhibit good convergence and stability of recurrent neural networks and show that the RNN is readily extendable to complex industrial problem.

Key words recurrent neural network; Elman network; back-propagation network; reaction process