

利用串/并联神经网络建立化学反应器模型

许兴华 曹柳林

(北京化工大学信息科学与技术学院, 北京 100029)

摘要: 研究了使用线性动态神经网络与非线性的静态网络相结合的混合建模方式解决复杂非线性系统的建模问题。使用混合神经网络建模,可以降低单个网络的训练难度,基于此,也可将非线性系统控制策略的求解分解,转换为线性系统的求解。从而改善使用单一神经网络建模存在的精度不高以及训练时间长等不足,也为非线性系统控制策略的求解提供方便。本文以一个典型多变量系统——连续搅拌釜式反应器(CSTR)作为仿真对象,详细研究和实现了两类神经网络串联和并联的混合建模方法,并对结果进行了比较。

关键词: B 样条网络; 线性递归神经网络; 混合神经网络; 连续搅拌釜式反应器(CSTR)

中图分类号: TP273.5

引言

全连接的递归神经网络,由于自身具有反馈结构,因而可以描述系统的动力学特性,很早就被用来建立动态系统的模型^[1]。为了模拟系统的非线性特性,网络必须包含非线性激励函数,常使用最优化方法,如梯度法,承担网络权值的训练。由于训练参数过多,训练过程收敛很慢,还存在可能陷入局部最小的通病。而 ELMAN、对角递归神经网络(DRNN)等其他一些本质动态的神经网络减少了网络连接参数,但常常要辅以输出变量的延时反馈,以弥补动态映射能力的不足^[2]。

本文采用混合建模的方式,将非线性静态网络与线性动态网络结合,分别承担模拟复杂非线性系统的相应部分特性。文章重点研究了两种网络的组合方式以及输出信号的综合模式,采用神经网络串联和并联两种混合方式,建立一个多变量连续搅拌釜式反应器(CSTR)系统的模型。文中的非线性静态网络采用了 BSNN 中的非线性映射部分,这部分的网络只需要计算 B 样条的基函数,不需要修正权值,从而降低了混合网络的训练难度;线性动态网络则采用了全连接的线性递归网络,有效避免了非线性网络在训练过程

中可能会出现局部极值等问题。

1 神经网络串/并联拓扑结构

文中的非线性静态网络采用了 B 样条网络(BSNN)中的非线性映射部分,这部分的网络只需要计算 B 样条的基函数^[3-4],不需要修正权值,从而降低了混合网络的训练难度;线性动态网络则采用了全连接的线性递归网络,有效避免了非线性网络在训练过程中可能会出现局部极值等问题。

由于当输入变量较多时,采用文献[5]中的多输入 B 样条网络基函数的算法,将会使网络的结构极其庞大,因此,本文采用(1)式计算基函数:

$$b_j(U) = \prod_{i=1}^m b_j(u_i), j=1, 2, \dots, s \quad (1)$$

其中, u_i 为系统第 i 个输入变量, m 为网络输入变量的个数, $b_j(u_i)$ 为 BSNN 第 i 个输入的第 j 节点基函数值, U 为系统输入向量, $b_j(U)$ 为多输入 B 样条网络第 j 节点基函数值。 s 为 B 样条网络的隐含节点数。

1.1 串联混合网络

串联混合网络的结构如图 1。

根据结构图可推导出网络输出的数学表达式:

$$\begin{cases} X(k) = AX(k-1) + GH(k) \\ Y(k) = CX(k) \end{cases} \quad (2)$$

其中, $H = [b_1(U), b_2(U), \dots, b_s(U)]^T$; $G \in R^{n \times s}$, $X \in R^{n \times 1}$, $U \in R^m \times 1$, $Y \in R^p \times 1$, $A \in R^{n \times n}$, $C \in R^p \times n$ 。

对(2)式进行 Z 变换,得(3)式。

$$Y(z) = C(ZI - A)^{-1}GH(z) \quad (3)$$

收稿日期: 2005-10-18

基金项目: 北京市教育委员会重点学科建设共建项目(XK100100435)

第一作者: 女, 1980 年生, 硕士生

E-mail: xxh.123@126.com

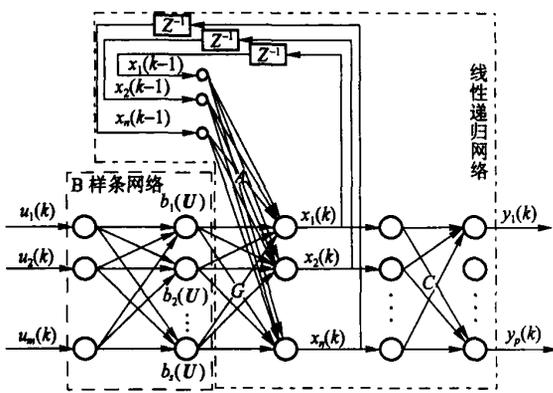


图 1 串联混合网络结构图

Fig.1 Structure of the serial hybrid neural network

由于 $(ZI - A)^{-1} = \frac{\text{adj}(ZI - A)}{\det(ZI - A)}$ ，其中， $\text{adj}(ZI - A)$ ， $\det(ZI - A)$ 分别为 $(ZI - A)$ 矩阵的伴随阵和行列式。

对式(3)进行整理，得(4)式。

$$Y(z) = \frac{\sum_{i=1}^p \sum_{j=0}^{n-1} \sum_{k=1}^s Q_{i,j,k} h_{i,j,k}(z) z^{-j}}{1 - \sum_{i=1}^n d_i Y(z) z^{-i}} \quad (4)$$

对(4)式进行反 Z 变换后得：

$$y_i(k) = \sum_{j=1}^n d_j y_i(k-j) + \sum_{j=0}^{n-1} \sum_{l=1}^s Q_{i,j,l} h_l(k-j) = \sum_{i=1}^T y_i(k)$$

采用递推最小二乘的方法求解模型参数 i_0 。

1.2 并联混合网络

并联混合网络的结构图如图 2。

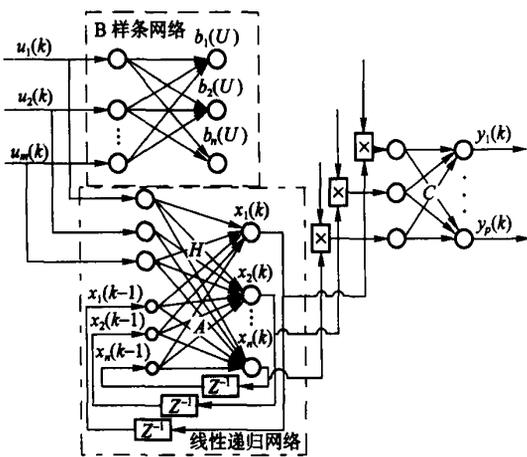


图 2 并联混合网络结构图

Fig.2 Structure of the parallel hybrid neural network

从结构图中，可知：B 样条网络节点个数与线性递归网络隐含层节点数相同。并联网络输出的数学表达式为

$$\begin{cases} X(k) = AX(k-1) + HU(k) \\ Y(k) = CB(U)X(k) \end{cases} \quad (5)$$

其中， $H \in R^{n \times m}$

$$B(U) = \begin{bmatrix} b_1(U) & 0 & \dots & 0 & 0 \\ 0 & b_2(U) & 0 & \dots & 0 \\ 0 & 0 & b_3(U) & \ddots & \dots \\ \dots & \dots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & b_n(U) \end{bmatrix}^{n \times n}$$

对(5)式进行 Z 变换，整理得(6)式。

$$Y(z) = CB(U)(ZI - A)^{-1}HU(z) \quad (6)$$

$$\text{由于 } (ZI - A)^{-1}H = \frac{D_i z^{-i}}{1 + \sum_{i=1}^n a_i z^{-i}}$$

$$D_i = \begin{bmatrix} d_{i,1,1} & d_{i,1,2} & \dots & \dots & \dots \\ d_{i,2,1} & d_{i,2,2} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ d_{i,n,1} & d_{i,n,2} & \dots & \dots & \dots \end{bmatrix}^{n \times m}$$

将(6)式整理为(7)式。

$$Y(z) = \frac{CB(U) \sum_{i=1}^{n-1} D_i z^{-i} U(z)}{1 + \sum_{i=1}^n a_i z^{-i}} \quad (7)$$

对(7)式进行反 Z 变换，并进行整理得：

$$y_r(k) = - \sum_{i=1}^n a_i y_r(k-i) + \sum_{i=0}^{n-1} \sum_{j=1}^m \sum_{l=1}^n b_l(U) Q_{i,j,l} u_j(k-i) = \sum_{r=1}^T y_r(k)$$

$$Q_{i,j,l} u_j(k-i) = \sum_{r=1}^T r_r(k)$$

其中， $r=1, 2, \dots$ 。采用递推最小二乘的算法，求解模型的参数 r_0 。

2 仿真研究

2.1 CSTR 模型描述

连续搅拌釜式反应器(CSTR)由(8)式描述^[6]：

$$\begin{aligned} \frac{dC_A}{dt} &= \frac{q}{V} (C_{Af} - C_A) - k_0 \exp\left(-\frac{E}{RT}\right) C_A \\ \frac{dT}{dt} &= \frac{q}{V} (T_f - T) + \frac{-H}{C_p} k_0 \exp\left(-\frac{E}{RT}\right) C_A + \frac{UA}{V C_p} (T_c - T) \end{aligned} \quad (8)$$

式中， C_A 是反应浓度； T 是反应温度； T_c 是冷却剂的温度； q 是反应物进料流速； C_{Af} 是进料浓度； T_f 是进料温度； V 是反应体积； k_0 是预指数因子； E 是活化能； H 是反应热； C_p 是热容。定义 q 和 T_c 为系统的输入，即 $U = [u_1 \ u_2]^T = [q \ T_c]^T$ ； C_A 和 T 定义

为系统的输出,即 $Y = [y_1 \ y_2]^T = [C_A \ T]^T$ 。在标称状态下,系统的参数见表 1。

表 1 系统参数表

Table 1 Parameters for the CSTR

C_A	C_{Af}	$-H$	T	UA	
0.2 mol/L	1 mol/L	17835.821J/mol	446 K	1000 g/L	11950J/(min·K)
V	C_p	T_f	E/R	K_0	q
100L	0.239J/(g·K)	400 K	5360 K	$\exp(13.4) \text{ min}^{-1}$	100 L/min

2.2 网络训练与测试

离散化 CSTR 数学模型,并取采样时间为 0.2 min。输入向量 U 在标称状态下做随机变化,采集训练样本数据。随机提取两组容量为 400、200 的两组数据,分别作为混合网络的训练样本和测试样本。

串联混合神经网络具体结构为(2,7,12,2),即:BSNN 样条网络部分取 2 个输入、7 个隐含节点,线性递归网络部分取 12 个隐含、2 个输出节点。BSNN 基函数阶数设为 3。反应物浓度的网络的测试效果如图 3。

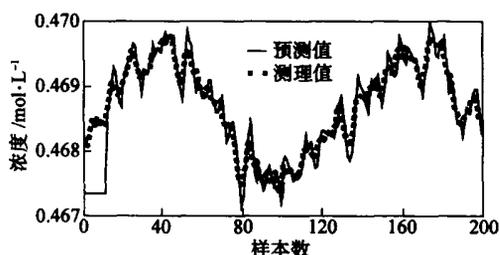


图 3 基于串联网路的测试结果

Fig. 3 Test results for the serial hybrid neural network

并联混合神经网络的结构为(2,7,7,2),即:BSNN 部分为 2 个输入、7 个隐含节点,线性递归网络部分设为 2 个输入、7 个隐含、2 个输出节点。B 样条网络基函数阶次取 3。相应的反应物浓度的网络的测试结果如图 4。

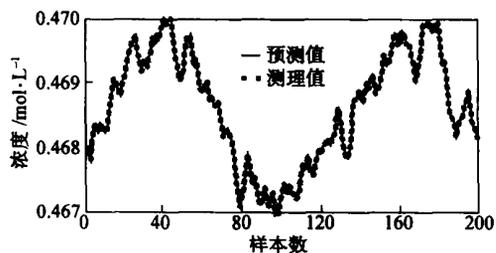


图 4 并联混合神经网络的测试结果

Fig. 4 Test results for the parallel hybrid neural network

通过仿真研究,可以证明串联与并联的混合神经网络均能够以很高的精度建立多变量 CSTR 对象的动态模型。

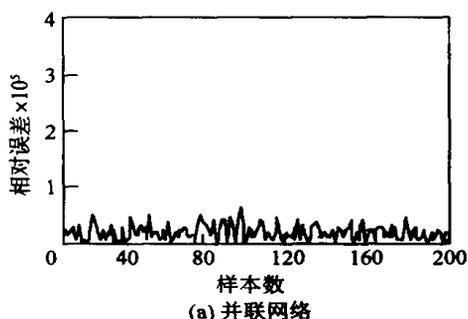
3 两种混合模型比较

3.1 训练参数比较

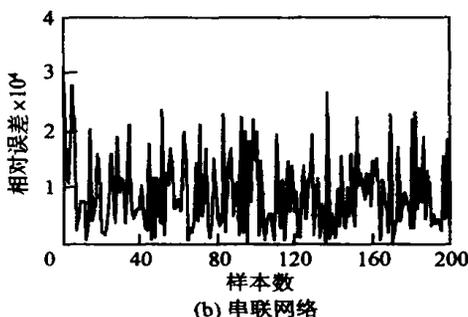
定义 M 为对象输入变量的维数, m 为 BSNN 部分隐含节点数, n 为线性递归网络部分隐含节点数, p 为输出节点个数。通过上述两种混合网络的结构描述可知,串联网路需要求解的参数为 $p \times m \times n + n$ 个。并联网路需要求解的参数为 $p \times M \times n^2 + n$ 个。当两混合网络采用相同的结构设置,均为 (M, n, n, p) ,BSNN 部分均采用 M 个输入节点, n 个隐含节点,线性递归网络部分取 n 个隐含节点, p 个输出节点。那么,并联网路需要求解的参数比串联网路多 $p \times (M - 1) \times n^2$ 个。

3.2 建模误差比较

两个网络测试结果的相对误差如图 5。



(a) 并联网路



(b) 串联网路

图 5 网络测试相对误差

Fig. 5 Test errors for the hybrid neural networks

本文采用公式(9)的平均相对误差(ARE)公式

$$\text{ARE} = \frac{1}{pL} \sum_{i=1}^p \sum_{k=1}^L \frac{|\hat{y}_i(k) - y_i(k)|}{y_i(k)} \quad (9)$$

式中, L 为测试数据的样本容量。 \hat{y} 为预测值, y 为测量值。

使用不同样本数量的数据训练和测试两个混合网络, 在达到相同精度前提下, 并联网络需要的训练数据样本数较少。设置两个网络结构相同, 均为 (2, 7, 7, 2), 使用相同的训练和测试样本对网络进行训练和测试, 计算得到并联网络测试的平均相对误差为 8.8177×10^{-6} , 而串联网络为 8.9387×10^{-5} , 前者比后者精度高多倍。

4 结束语

使用混合神经网络建模, 可以降低单个网络的训练难度, 同时基于此, 也可将非线性系统控制策略的求解分解, 转换为线性系统的求解。从而改善使用单一神经网络建模存在的精度不高以及训练时间长等不足, 也为非线性系统控制策略的求解提供方便。

本文通过使用 CSTR 作为仿真对象, 使用串联和并联两种混合网络分别建模, 证明使用混合网络可以

以较高的精度建立复杂非线性的模型。

实验结果表明在使用相同的网络结构时, 并联混合网络的建模精度比串联网络建模精度高, 但由于训练参数多, 训练时间略长。

参 考 文 献

- [1] Wang Deliang, Liu Xiaomei. On temporal generalization of simple recurrent networks [J]. *Neural Networks*, 1996, 9: 1099 - 1118.
- [2] 赵奇, 刘开第, 庞彦军. Elman 神经网络训练方法及其在非线系统辨识中的应用 [J]. *煤矿机械*, 2005, 5 (5): 73 - 74.
- [3] 曹柳林, 吴海燕. 利用 B 样条神经网络实现聚合反应分子量分布的建模与控制 [J]. *化工学报*, 2003, 55 (5): 742 - 746.
- [4] Wang H. Bounded dynamic stochastic distributions [M]. *Modeling and Control*, London: Springer-Verlag, March, 2000, 17 - 31.
- [5] 於东军, 王士同. B 样条神经网络的构造理论 [J]. *J Comp Res Dev*, 1999, 36: 534 - 540.
- [6] 周东华. 现代故障诊断与容错控制 [M]. 北京: 清华大学出版社, 2000.

Modeling of a chemical reactor with serial/ parallel neural networks

XU Xing-hua CAO Liu-lin

(College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: Modeling studies of complex nonlinear systems by hybrid neural networks are described in this paper. The hybrid neural networks consist of linear dynamic neural networks and nonlinear static neural networks. Using such a hybrid neural network, the difficulties in training a single network can be decreased, and the solution for a nonlinear control strategy can be reduced to solving for a linear system. Thus, this method overcomes the shortcomings of long training time and lower accuracy observed for one single neural network. Both the serial and parallel neural networks have been used to model a CSTR. A comparison between the two hybrid neural networks on the basis of the results obtained is described.

Key words: BSNN; linear recurrent neural network; hybrid neural network; CSTR