

神经网络过程模型辨识*

朱群雄 麻德贤

(北京化工大学计算机系, 北京 100029)

摘 要 利用修正的反传神经网络(BP)算法分别对稳态过程和动态过程以及两个实际化工生产过程进行模型结构辨识, 所确定的最佳神经网络结构经验证表明具有较好的收敛性和稳定性, 模拟精度高、适用性强。

关键词 神经网络 模型辨识 反应工程

引言

建立过程的数学模型是进行过程模拟与过程控制的必要基础。通常建模的方法主要是采用机理法和测试法, 但由于许多过程的复杂性和不确定性, 往往很难建立其机理模型, 也很难基于测试数据进行常规的模式识别。由于人工神经网络在解决高度非线性和严重不确定性系统的辨识与控制方面具有一系列优于其他计算方法的性质^[1], 因而近几年受到人们的普遍关注。本文将通过应用修正的BP神经网络算法对有关过程进行模型化与模拟, 以证明神经网络过程模型辨识的可靠性和实用性。

1 BP算法的改进

Hornik 等人^[2]证明了用一个隐含层和足够多的节点可以实现任意精度的输入-输出映射关系, 但是, 对结构确定目前尚没有理论上的指导, 人们大多是基于经验或某些原则人为确定网络结构, 如依据网络输出误差的大小逐步增加或减少节点, 显然, 这种方法要花费很多时间, 而且很难得到最佳结构。一般认为, 大的网络结构容易提高训练精度, 但也容易使之失去整个过程的特性而不能对训练样本以外的数据进行很好的预测。因此, 网络结构的确定直接影响网络映射精度、稳定性和适应性。然而, 从神经网络的原理可知, 神经网络是通过节点间的权值来传递信息的, 如果权值对输入信息既不起激励作用、也不起抑制作用, 那么, 这个权值就是冗余的, 就可以去掉这个权值对应的节点间的连接线。对于某个节点, 如果与上一层节点的所有连接权值都是冗余的, 或者该节点与下一层所有节点的连接权值都是冗余的, 那么, 这个节点也就是冗余的, 可以删除之。由此可知, 通过辨别和删除冗余权值并进而辨别和删除冗余节点可以实现网络结构的简化。权值的正负作用是与节点上的活化函数有关, 如果采用以零点为中心的活化函数, 那么, 当权值为正时, 表明该权值对网络起激励作用, 权值为负时则对网络起抑制作用, 而权值为零时则对网络不起作用, 该权值就是冗余的, 可以删除。为此, 可以对标准Sigmoid函数进行改进, 新的以零点为中心的网络节点

1996-12-13 收到初稿, 1997-05-09 收到修改稿。

联系人及第一作者: 朱群雄, 男, 37岁, 博士, 副教授。

* 国家自然科学基金资助项目 (No. 29676002)。

活化函数如下式所示

$$f(x) = \frac{1}{1 + e^{-x}} - 0.5 \quad (1)$$

当一个网络训练至稳定收敛后,所有具有零值的权值连线可以断开,进而可以删除冗余节点.但是,由神经网络计算原理可知,网络在训练过程中对每一个权值的修正策略是一样的,并不能使冗余权值自动“暴露”出来或逼使其值趋于零.因此,采用这种以零点为中心的活化函数只是具备了辨别冗余权值和冗余节点的条件,为迫使冗余权值自动趋于零值和使网络结构简化,本文采用双重优化目标函数,即在优化网络中各权值使网络输出误差函数 E_{BPN} 最小化的同时,对冗余权值进行惩罚使之变为零值而被删除,以减少网络节点和权值数使体现网络结构大小的函数 E_{STR} 最简化,如下式所示

$$\min E_T(w, N_w) = E_{\text{BPN}} + \lambda E_{\text{STR}} \quad (2)$$

$$E_{\text{BPN}} = \frac{1}{2} \sum (y_j - v_j)^2$$

$$E_{\text{STR}} = \sum_j N_j \sum_{i=1}^{N_j} w_{ij}^2$$

式中, w 为一组网络权值; N_w 是网络中权值总数; N_j 为进入节点 j 的权值个数; λ 是一个待定系数.显然,从某种意义上说, E_{BPN} 和 E_{STR} 是矛盾的一对函数.只有使 E_{BPN} 和 E_{STR} 处于最佳协调关系中,使总体目标函数最优,才可认为最佳网络结构找到了.利用非线性规划中的最速下降法,由 $\Delta w_{ij} \propto -\frac{\partial E_T}{\partial w_{ij}}$ 就可以得到网络权值修正公式如下

$$\Delta w_{ij} = \eta \delta_j^k v_i^{k-1} - \gamma N_j w_{ij} \quad (3)$$

输出层的 δ_j^k 计算公式为

$$\delta_j^k = (y_j - v_j^k)[0.25 - (v_j^k)^2] \quad (4)$$

隐含层的 δ_j^k 计算公式为

$$\delta_j^k = [0.25 - (v_j^k)^2] \sum_i w_{ji} \delta_i^{k+1} \quad (5)$$

这里的 η 是学习速率, $\gamma = 2\eta\lambda$. 是个十分敏感的参数.本文提出如下自适应计算公式

$$\gamma = \gamma_0 e^{-E_{\text{BPN}}} e^{-|w_{ij}|} \quad (6)$$

这样, γ 就依据 E_{BPN} 和权值的大小进行自动调整. γ_0 是一个需人为给定的初始值,由于 γ 具有较好的自适应能力,本文建议 γ_0 一般可取 0.0001.

神经网络结构辨识的算法计算步骤是:先根据对象初设一个较大结构的网络和收敛精度,运行本算法,当网络在收敛区时就可自动地逼使冗余权值趋于零而被删除,进而可以辨别和删除冗余节点.由此,神经网络就可得以简化.值得注意的是,绝对的计算零值是少有的,因此,可设一个权值死区,当权值很小时,就可认为它是冗余的.

2 稳态过程的神经网络模型结构辨识

在实际神经网络过程模型辨识中,往往除了要对隐含层上的节点进行辨识外,由于对影响过程输出的输入变量难于明确,也要对输入的参数进行辨识.作者以下面的一个具体的函数关系来确定其神经网络结构.

$$x_1 = 0.5x + 4.57$$

$$x_2 = 2.5x + 1.23$$

$$x_3 = \sin[\text{rand}() / 360.0] \times 5.0$$

$$y = x_1 + x_2 \quad (7)$$

显然, 这里的 x_3 是一个冗余输入变量. 利用 C 语言中的随机数发生器产生三入一出的 40 组训练数据和 40 组测试数据, 分别对隐含层节点数为 20、10、7、5 等四种情况的三层神经网络进行结构优化. $\gamma_0 = 0.0001$, 权值死区定义为 ± 0.05 . 当四种网络结构学习次数分别为 3000、1816、658 和 1231 时删除了冗余输入 x_3 ; 当学习次数分别为 3647、2677、671 和 1290 时, 四种情况下的网络结构均为 2 入 2 中 1 出, 再分别往下继续训练, 四种网络结构都保持不变了. 需要说明的是, 尽管四种结构中隐含层上的两个节点位置不同, 但其物理意义是相同的. 图 1 给出了四种情况下的网络结构辨识过程. 此外, 再分别用 2 入 1 中 1 出和 2 入 3 中 1 出建立相应的神经网络模型, 其最后的训练和模拟精度都不如 2 入 2 中 1 出的模型好. 这个例子说明本文所提出的结构学习算法既能删除隐含层中的冗余节点, 也能删除输入层中的冗余变量, 所得到的模型结构是最佳的.

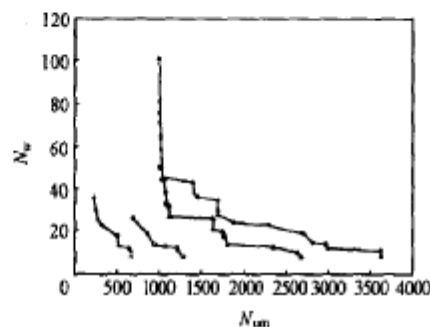


Fig.1 Structure recognition for the four networks
structure: * 1; Δ 2; + 3; $\circ$ 4

3 动态过程的神经网络模型结构辨识

连续搅拌釜反应器 CSTR(Continuous Stirred Tank Reactor)是一个公认的、典型的化工动态过程实例. 文献[3]描述了一个不可逆的放热化学反应: $A \rightarrow B$. 这是一个温度随时间变化的非线性系统, 要通过调节冷却水流量来控制组分 A 的浓度 C_A . 这里, 要根据过程的采样数据来建立组分浓度 C_A 与冷却水流量 q 之间的时间离散化的动态关系模型, 基于神经网络的动态过程模型可用下式表示

$$C_A(t+1) = NN[w, C_A(t), C_A(t-1), \dots, C_A(t-n), q(t), q(t-1), \dots, q(t-n)] \quad (8)$$

式中的 NN 是神经网络非线性函数.

取采样周期为 0.1 min, 随机获取一系列开环数据. 先取当前时刻及前二个时刻的冷却水流量 $q(t)$, $q(t-1)$, $q(t-2)$ 和反应器中的组分 A 浓度 $C_A(t)$, $C_A(t-1)$, $C_A(t-2)$ 作为神经网络的输入变量, 下一个时刻的浓度 $C_A(t+1)$ 作为神经网络的输出, 随机产生这样的数据 200 组用于训练神经网络, 再产生 100 组数据用于测试网络. 初选一个隐含层并设有六个节点, 因此, 初始神经网络为 6-6-1 结构, 其收敛结果如表 1 所示. 利用本文提出的算法对该神经网络进行结构优化计算, 隐含层节点数最后从六个下降到二个, 输入层中冷却水流量即反应过程的控制变量只取了二个, 删除了当前时刻的控制变量, 这就意味着当前时刻的控制作用对下个时刻的被控对象不起作用或作用甚微而被忽略, 这是符合这个对象的动态特性的, 因为冷却水流量的调节对反应器中的反应温度有影响, 进而控制反应浓度, 这是

一个滞后的过程. 对作为输入变量的反应浓度 C_A 仍保留了三个时刻的变量, 说明前三个时刻的反应浓度对下个时刻的反应有明显的影响. 网络权值总数从初始的 49 个下降为九个, 且隐含层和输出层均没有阈值.

为了考察这个网络结构是否是最佳的, 即是否遗漏其他时刻的变量对过程的影响, 按同样的方法随机仿真产生了各具四个时刻的冷却水流量和反应浓度作为输入变量和下个时刻的反应浓度作为输出变量的两个数据文件用于训练和测试神经网络, 数据文件分别有 200 组和 100 组数据. 这些输入变量是 $q(t)$, $q(t-1)$, $q(t-2)$, $q(t-3)$ 和 $C_A(t)$, $C_A(t-1)$, $C_A(t-2)$, $C_A(t-3)$, 输出变量是 $C_A(t+1)$. 取一个隐含层并仍设六个节点, 这样, 初始神经网络为 8-6-1 结构. 经运行结构优化算法, 网络权值总数从 61 个下降到九个, 最终网络结构仍为 5-2-1, 且所选的输入变量与上一种情况的结果是相同的. 表 1 给出了三种不同网络结构下的收敛情况, 由此可以看出结构 5-2-1 为最佳, 证明了本文的方法是可靠的.

Table 1 Training results on three neural networks

Network structure	Iteration	Training set			Test set		
		Output MSE	Relative error	Absolute error	Output MSE	Relative error	Absolute error
network 1 6-6-1	700	0.000354	0.7083	0.0006	0.000143	0.3669	0.0003
network 2 8-6-1	1000	0.000292	0.9305	0.0008	0.000143	0.3492	0.0003
network 3 5-2-1	2000	0.000139	0.2870	0.0002	0.000056	0.2129	0.0002

4 工业过程模型化与模拟的应用

4.1 萘催化氧化流化床反应器神经网络模型

本文研究对象是 $\Phi 3.6\text{ m}$ 的垂直式内构件萘催化氧化流化床反应器, 用 V_2O_5 作催化剂, 其氧化产物有邻苯二甲酸酐(简称苯酐)及其他多种副产物. 影响苯酐选择性的因素很多, 主要可操作参数是反应温度、空萘比和送风量等^[4]. 由于该反应机理和流化床模型极其复杂, 建立能反映实际生产装置的数学模型是相当困难的. 因此, 根据工业现场要求选择反应温度、空萘比和送风量三个参数作为神经网络的输入变量, 苯酐的选择性作为输出变量, 从生产技术台账和现场实地取样分析的数据中整理出 70 组用于训练神经网络和 30 组用于测试网络. 初设网络结构为 3-10-1, 利用本文提出的新算法对此进行了结构辨识和简化, 最后得到 3-4-1 的结构, 基于这个结构的工业模拟结果如图 2 和图 3 所示. 由此可知, 所建神经网络模型具有较好的模拟精度和较好的工业应用适用性.

4.2 苯乙烯生产装置烃化/反烃化反应器的神经网络模型

苯乙烯装置乙苯单元的主要设备是进行烃化/反烃化反应的套筒式烃化反应器, 乙烯和苯在催化剂的作用下在反应器的内筒进行烃化反应^[5], 反应液上升到反应器内筒的上部并溢流到反应器的外筒, 与精馏区来的回流多乙苯进行反烃化反应, 最后反应液流出反应器送到后续装置进行处理. 提高反应液中的乙苯质量百分含量是提高乙苯生产产量、降低乙烯和苯原料消耗的主要目标, 因此, 以烃化液中乙苯的质量百分含量作为神经网络的输出变量. 根

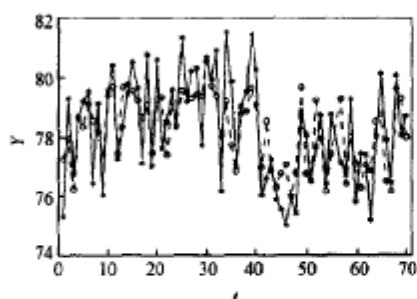


Fig.2 Prediction of training data

average absolute error is 1.1177;

average relative error is 1.4309%;

* actual data; ○ network simulation

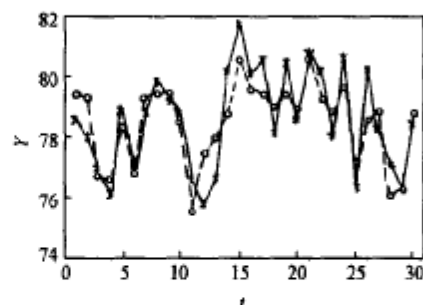


Fig.3 Prediction of test data

average absolute error is 1.1359;

average relative error is 1.4479%;

* actual data; ○ network simulation

据工业生产分析,影响目标变量的可操作参数主要是乙烯进料量、乙烯与苯的摩尔比、反应温度、多乙苯回流量和催化剂加入量等,因此,选择这五个参数作为神经网络的输入变量.经对现场生产数据严格筛选,整理了 60 组数据用于训练网络和 35 组数据用于测试网络.网络的初始结构为 5-10-1,对这个网络进行结构优化计算后,最后得到 5-6-1 的模型结构.基于工业生产数据的模拟结果如图 4 和图 5 所示.结果表明,神经网络结构优化技术是可靠的,神经网络过程模型辨识具有较好的工业实际应用前景.

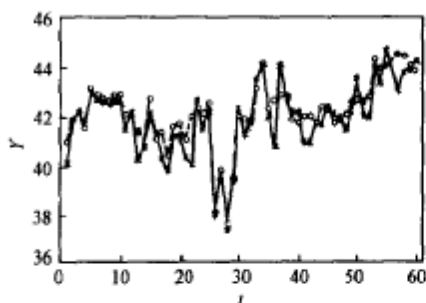


Fig.4 Prediction of training data

average absolute error is 0.5148;

average relative error is 1.2287%;

* actual data; ○ network simulation

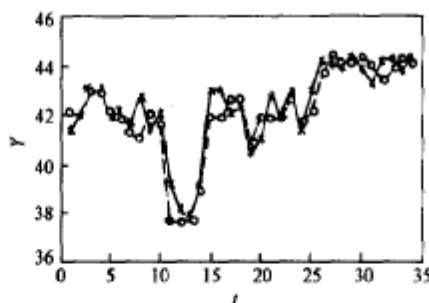


Fig.5 Prediction of test data

average absolute error is 0.5889;

average relative error is 1.3968%;

* actual data; ○ network simulation

5 结论

从上面稳态、动态和实际生产过程等四个实例的模型化和模拟情况,可以看出所用的基于 BP 算法的神经网络结构优化计算方法具有较好的适应性,所建模型结构简单、模拟精度高.因此,通过神经网络结构辨识可以得到符合实际过程的模型并以此进行较好的过程模拟,本法可在许多场合的过程建模中应用.

符 号 说 明

C_A ——组分 A 浓度, $\text{mol} \cdot \text{L}^{-1}$	q ——冷却水流量, $\text{m}^3 \cdot \text{s}^{-1}$
E_{BPN} ——网络输出误差函数	t ——时间序列
E_{STR} ——网络结构函数	v_j ——输出层第 j 个节点的网络输出
E_T ——总的网络目标函数	Δw_{ij} ——节点 i 与节点 j 间的权值更新量
MSE ——网络输出均方差	x ——变量
N_j ——进入节点 j 的权值个数	Y ——对象目标输出
N_{un} ——训练迭代次数	y_j ——输出层第 j 个节点的期望输出
N_w ——网络中权值总数	γ ——网络结构调整参数
	η ——学习速率

参 考 文 献

- 1 查金荣, 申同贺, 包宏等译. 人工智能在化学工程中的应用. 北京: 中国石化出版社, 1994. 450
- 2 Hornik K, Stinchcombe M, White H. Neural Networks, 1990, 2:359 ~ 366
- 3 Henson M A, Seborg D E. AIChE JI, 1990, 36:1753 ~ 1757
- 4 Zhu Qunxiong, Kan Danfeng, Ma Dexian. Computers Chemistry Monograph Series 2. Beijing: Science Press, 1991. 162 ~ 165
- 5 于成烈, 朱群雄, 麻德贤等. 合成树脂与塑料, 1994, 11(2):38 ~ 42

RECOGNITION OF NEURAL NETWORK PROCESS MODEL

Zhu Qunxiong and Ma Dexian

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

Abstract A novel algorithm of neural network structure optimization is described in this paper. The capabilities of the algorithm are demonstrated by application in steady-state process, dynamic process and industrial process. The recognized results of neural network structure show that the optimal neural network can be widely used for modelling and simulation of chemical processes.

Keywords neural network, model recognition, reaction engineering