

一种利用多神经网络结构 建立非线性软测量模型的方法*

熊智华 王 雄 徐用懋
(清华大学自动化系 北京 100084)

摘 要 采用多神经网络结构建立软测量模型。建模所用的样本数据经过 Kmeans 聚类方法分成多组训练数据, 每组数据建立一个单个神经网络模型, 再通过 PCR 方法连接起来得到整个模型的输出, 从而显著地提高了模型的精确度和鲁棒性。仿真研究表明, 采用该建模方法能够达到较好的建模效果。

关键词 多神经网络, 软测量模型, Kmeans 聚类, PCR 方法

分类号 TP 216

Nonlinear Software Sensor Modeling Using Multiple Neural Network

Xiong Zhihua, Wang Xiong, Xu Yongmao
(Tsinghua University)

Abstract Software sensor has been considered as an efficient tool to deal with the difficulty of unmeasurable primary output. Generalization of neural network model can be significantly improved by combining several models. The approach to build software sensor model by multiple neural networks (MNN) based on Kmeans clustering method is presented. The MNN model is developed by several feedforward neural network models. Those models are combined by principal components regression. Software sensor model accuracy is perfectly improved. The simulation results demonstrate efficiency of the method.

Key words multiple neural network, software sensor, Kmeans clustering, principal components regression

1 引 言

在许多工业过程控制系统中, 软测量技术较好地解决了输出变量难以测量的问题, 因而在理论研究和工业应用中得到了广泛发展^[1,2]。近年来, 将神经网络引入软测量技术的建模方法已引起许多学者的研究兴趣^[3,4]。对于非线性对象采用基于神经网络的软测量建模方法, 要解决的主要问题是神经网络模型的鲁棒性和模型的预测精确度。目前已有多种神经网络训练方法用来增强网络的鲁棒性和精确性, 其中通过建立一组神经网络模型, 并按一定方式将其连接起来以改善模型预测能力的方法受到许多

学者的关注^[5-7], 并提出各种网络结构, 例如 Multiple Neural Networks (MNN), Stacked NN (SNN) 和 Distributed NN (DNN) 等。其中 MNN 模型结构就是这种方法的基本组成形式, 不同结构之间的区别在于不同的连接方法。

本文提出一种基于 MNN 建立软测量模型的方法。建立神经网络模型的样本数据集通过 Kmeans 聚类方法分成多组训练和测试数据, 每组训练和测试数据建立一个单个神经网络模型, 各单个神经网络模型的预测输出通过主元递归 (PCR) 方法连接, 从而得到整个模型的输出。PCR 是一种多元投影方法, 能够较好地解决数据相关的问题, 已在统计分析方法中得到广泛应用。由于各单个神经网络模型反映同一种非线性关系, 因此它们是严重关联的, 采用

* 1998- 10- 12 收稿, 1998- 12- 31 修回

PCR 方法可以得到鲁棒的连接权系数。仿真研究表明,该方法能够显著提高模型的精确度和鲁棒性。

2 基于 Kmeans 聚类方法的多神经网络模型

连接多个模型以改进模型预测能力的方法是由 Bates^[8]于 1969 年提出的。该方法在时间序列分析中得到较广泛的研究,近年来在神经网络的研究中也备受关注^[5-7]。当用系统输入输出数据建立非线性对象的神经网络模型时,采用单个神经网络建立的模型往往只是系统的一种近似模型^[9],而且不同网络在不同输入空间中的预测性能会有所不同。而多个神经网络通过一定方式将这些单个网络进行连接,构成对象的整个输入空间模型,模型的预测精确度和鲁棒性得到了增强。连接多个网络模型的一般化结构是模块式泛化结构。

基于 Kmeans 聚类方法的多神经网络(MNN)模型结构如图 1 所示,它是由多个相对独立、协同作用的单个神经网络组成的系统^[5]。对象的输入空间通过 Kmeans 聚类等方法划分为一些小的局部空间,在每个局部空间中用单个神经网络实现期望的非线性映射关系,或使某一单个网络在此局部空间中发挥主要作用,在整体上由 MNN 模型实现整个输入空间所期望的函数映射关系。单个网络的结构形式则可选取具有良好非线性映射的逼近能力的网络,如 MLP、RBF 网络等。

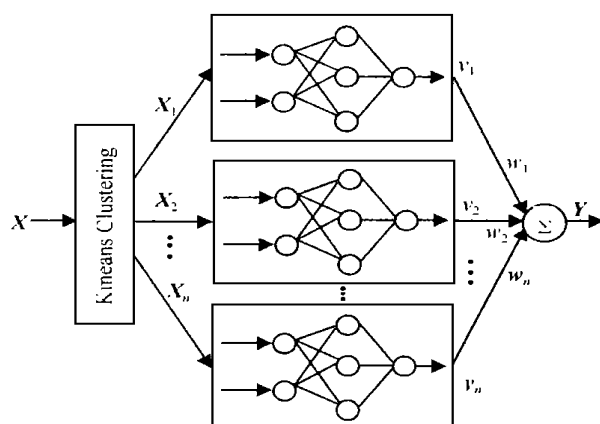


图 1 基于 Kmeans 聚类方法的多神经网络结构

MNN 模型的输出是所有单个神经网络输出的连接,可表示为

$$Y = \sum_{i=1}^n w_i y_i = \sum_{i=1}^n w_i f(X_i) \quad (1)$$

其中, Y 是 MNN 预测输出, y_i 是第 i 个神经网络的预测输出, w_i 是连接第 i 个神经网络的权值, X_i 是经过 Kmeans 聚类方法得到的第 i 个神经网络的输入向量。

3 用 PCR 方法计算 MNN 模型的连接权值

MNN 模型中计算连接权值的方法有等权值法,最小二乘法 and 模糊分类法^[7]等。Dasaratha^[5]采用 RMSE (误差均值平方根) 指标,对所训练的多个神经网络进行取舍后构成 SNN 模型,其中模型连接权值的计算采用最小二乘法。这种方法的不足之处是,由于所训练的多个神经网络模型反映的是同一种非线性关系,其相互之间是严重相关的,采用最小二乘法难以得到较好的连接权值,而且 SNN 方法舍弃了经过训练的神经网络,从而丢失了这些模型中含有的有用信息,使整个模型的预测精度下降。

针对 SNN 模型中连接权值计算方法的不足,可采用 PCR 方法来综合多个神经网络模型的预报值,解决其相互之间的严重相关问题,从而得到鲁棒的连接权值。因此,本文采用 PCR 方法来计算 MNN 中的连接权值。

考虑对象为多输入单输出(MISO)的情况。假定 MNN 模型由 n 个单个神经网络构成, Y_{sp} 是模型输出期望值, y_i 是第 i 个神经网络模型的预测输出值。将 n 个预测输出值写成向量形式为

$$\hat{Y} = [y_1 \ y_2 \ \cdots \ y_n]^T$$

以 $\hat{y}_{MNN}$ 表示 MNN 模型的预测值,则由式(1)可知

$$\hat{y}_{MNN} = W^T \hat{Y} = w_1 y_1 + w_2 y_2 + \cdots + w_n y_n \quad (2)$$

其中 $W = [w_1 \ w_2 \ \cdots \ w_n]^T$ 为连接权值向量。

对于 m 组样本数据, n 个单个网络可以得到 m 组预测值 $\hat{Y}$, 将其构成矩阵 H ($m \times n$), 并将与其对应的 m 组输出期望值构成向量 Y_{sp} ($m \times 1$)。求解参数 W 的传统方法是采用最小二乘法,即

$$W = H^+ Y_{sp}$$

其中 $H^+ = (H^T H)^{-1} H^T$, 可以通过求解 H 的 SVD 来得到。

采用 PCR 方法将 H 写成另一种形式,即

$$H = t_1 p_1^T + t_2 p_2^T + \cdots + t_n p_n^T \quad (3)$$

式中, $t_i = u_i \sigma_i$, $p_i = v_i$, 分别称为第 i 个主元的主元分量和负荷分量,均为单位正交向量,而且 $t_i = H p_i$ 。一般取反映对象主要变化部分的前 k 项来计算,从而有

$$H \approx H_k = TP^T \quad (4)$$

其中, $T = [t_1, t_2, \dots, t_k]$, $P = [p_1, p_2, \dots, p_k]$ 。则由式(2)可得

$$Y_{sp} = HW = TP^T W \quad (5)$$

记 $W_k = P^T W$ 。由上式可计算 W_k 的最小二乘解为

$$W_k = (T^T T)^{-1} T^T Y_{sp} \quad (6)$$

由于 P 为正交向量, 即有 $P^T = P^{-1}$, 从而可得连接权值 W 为

$$W = PW_k = P(T^T T)^{-1} T^T Y_{sp} \quad (7)$$

4 非线性对象的软测量模型仿真研究

4.1 复杂非线性对象的MNN建模方法

考虑如下非线性对象的建模问题

$$y = 10 \sin(\pi x_1 x_2) + 20(x_3 - 0.5) + 10x_4 + 5x_5 + x_6 \quad (8)$$

式中 $x_i (i = 1, 2, \dots, 6)$ 服从均值为 0, 方差为 1.0 的均匀分布。Dasaratha^[5] 亦采用该非线性函数来分析 SNN 建模方法。

建模所用的数据由下列方法得到: 训练样本集 S_1 包含 200 个样本, 样本中加入 SNR 为 4.0 的高斯噪声, 其中 $SNR = \sigma_y / \sigma_n$, σ_y 为输出的标准偏差, σ_n 为噪声的标准偏差。另外, 还生成不含噪声的 100 个样本数据构成测试数据集 S_2 , 用来检验不同网络模型的性能。训练样本集 S_1 采用 Kmeans 聚类方法进行聚类, 得到 4 个极值中心, 进而将样本数据分成 4 部分。对于每一部分的数据建立单神经网络模型, 其中采用 RMSE 来衡量网络模型的性能。网络结构采用由 10 个隐层节点构成的两个结构相同的单隐层前向网络模型 N_1, N_2 , 以及由 20 个隐层节点构成的单隐层前向网络模型 N_3, N_4 。其中相同的网络采用不同的初始权值和训练停止依据。网络训练算法采用 Levenberg - Marquardt 算法, 这是一种改进的 BP 算法, 具有训练速度快的特点。在确定上述 4 个网络后, 将其按 PCR 方法进行连接, 得到如下 MNN 模型

$$y_{MNN} = 0.304y_{N_1} + 0.352y_{N_2} + 0.142y_{N_3} + 0.105y_{N_4} \quad (9)$$

利用测试集 S_2 计算 RMSE 值和网络的最大绝对预测误差 MaxError 来衡量各网络模型的性能。分别计算 4 个 M FNN 网络以及 MNN 模型的 RMSE 和 MaxError 值, 表 1 给出了各个网络性能的比较结

果。从表 1 可以看出, 由于初始权值不同以及训练停止依据不同, 相同的网络结构其性能会有一定差别, 而 MNN 的性能也相应地得到很大改善。

表 1 模型的 RMSE 和 MaxError 比较

Model	RMSE		MaxError	
	S_1	S_2	S_1	S_2
$N_1: 6-10-1$	0.026 0	0.040 1	0.110 8	0.328 0
$N_2: 6-10-1$	0.023 3	0.037 1	0.129 8	0.267 9
$N_3: 6-20-1$	0.040 2	0.060 6	0.179 3	0.451 3
$N_4: 6-20-1$	0.068 0	0.095 2	0.223 4	0.517 4
MNN	0.020 2	0.030 8	0.082 7	0.215 7
SNN	0.031 8	0.045 4	0.134 5	0.263 5

将基于 PCR 综合方法的 MNN 模型与 SNN 建模方法进行比较。按照 SNN 模型的确定方法^[5], 采用上述 4 个模型中的 N_1, N_2, N_3, N_4 个模型, 利用最小二乘法计算 SNN 模型的连接权值, 得到

$$y_{SNN} = 0.352y_{N_1} + 0.386y_{N_2} + 0.261y_{N_3} \quad (10)$$

同样计算 SNN 模型的 RMSE 和 MaxError 值。从表 1 可以看出, 基于 PCR 综合方法的 MNN 模型在模型准确性方面比 SNN 建模方法效果好。

4.2 精馏塔产品质量软测量模型

丙烯精制精馏塔是一种典型的工业精密精馏塔, 对塔顶的丙烯产品质量要求很高, 达到 99.5%, 因而要求很高的控制精度。但产品质量难以提供实时信号, 因此建立丙烯质量的软测量模型, 可为实现质量闭环控制, 提高控制精度提供一种可靠方法。

建立丙烯质量的软测量模型需要选择好二次变量。根据对丙烯精馏塔的工艺流程和控制系统结构的了解和分析, 选取二次变量为塔顶的温差 ΔT 、双温差 $\Delta^2 T$ 、回流量变化 ΔR 以及塔顶压力变化 ΔP ^[10]。软测量模型的一般表达式为

$$y = f(\Delta T, \Delta^2 T, \Delta R, \Delta P) \quad (11)$$

以某厂丙烯塔的非平衡级稳态机理模型为对象进行分析^[11], 获得了建立软测量模型所需的样本数据, 其中包含了回流量变化、压力变化以及回流量和压力同时变化 3 种情况下的样本。经过归一化处理后得到 97 个样本数据, 通过随机抽取的方法生成含 57 个样本的训练样本集 S_1 和含 40 个样本的测试样本集 S_2 。训练样本数据集 S_1 通过 Kmeans 聚类方法进行聚类后, 得到 3 个极值中心, 进而将样本数据分成了 3 部分。网络结构选用由 10, 15, 20 个隐层节点构成的单隐层多层前向网络模型 N_1, N_2, N_3 。网络的训练算法采用 Levenberg - Marquardt 算法,

然后采用 PCR 方法将上述 3 个网络连接得到如下 MNN 模型

$$y_{MNN} = 0.318y_{N_1} + 0.816y_{N_2} + 0.273y_{N_3} \quad (12)$$

同样采用测试数据集 S_2 的 RMSE 和 MaxError 值来衡量各网络模型的性能。

表 2 给出了各个网络的性能比较结果。可以看出, MNN 模型的 RMSE 减少了许多, MaxError 值也相应地降低了, 表明其较好地改进了模型的性能。

表 2 模型的 RMSE 和 MaxError 比较

NN Model	RMSE	MaxError
$N_1: 4-10-1$	0.064 4	0.013 0
$N_2: 4-15-1$	0.081 5	0.019 1
$N_3: 4-20-1$	0.081 1	0.018 9
MNN	0.053 8	0.011 7
SNN	0.073 5	0.016 8

采用上述 3 个模型来构成 SNN 模型。计算连接权值后得到的模型为

$$y_{SNN} = 0.445y_{N_1} + 0.243y_{N_2} + 0.312y_{N_3} \quad (13)$$

并计算了 SNN 模型的 RMSE 和 MaxError 值。从表 2 可以看出, 基于 PCR 综合方法的 MNN 模型在模型准确性方面比 SNN 建模方法效果好。图 2 为 MNN 模型和 SNN 模型的预报结果比较(“+”为非线性模型, “○”为 MNN 模型, “▽”为 SNN 模型), 也证明基于 PCR 方法的 MNN 模型在模型的预报能力上有了一定的改善。

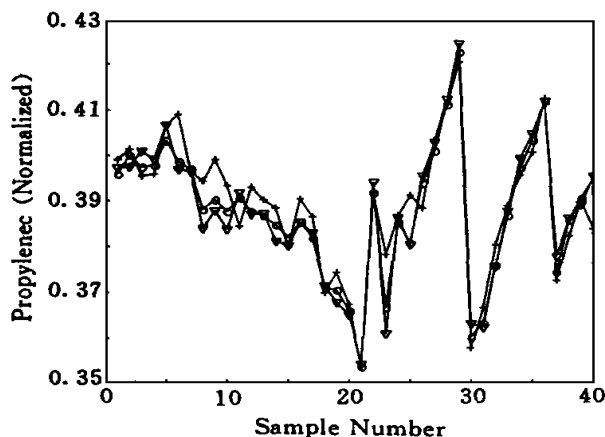


图 2 MNN 模型与 SNN 模型比较

5 结 论

本文将 MNN 模型应用于建立非线性对象的软测量模型, 提出了采用 PCR 方法综合的 MNN 建模方法, 解决了多个单神经网络模型间的严重关

联问题, 增强了模型的预测能力, 提高了模型的鲁棒性。仿真研究表明, 采用该建模方法能够达到较好的建模效果, 并能有效地改善软测量模型的预测能力及推广能力。但是, 采用 MNN 建立软测量模型的方法还需要在工业过程的实际应用中进一步完善和深入研究。

参 考 文 献

- 1 Bhat N V, M indeman P A, M cA voy T J *et al*. Modeling chemical process systems via neural computation. IEEE Control Systems Magazine, 1990, April: 24- 29
- 2 Joseph B, Brosilow C B. Inferential control of processes. A IChE Journal, 1978, 24(3): 485- 508
- 3 Andreas D, Sebastian E, Horst R. Model predictive control using neural networks. IEEE Control Systems Magazine, 1995, October: 61- 66
- 4 Willis M J, Massimo C D, Morris A J *et al*. Artificial neural networks in process engineering. IEE Proc D, 1991, 138(3): 256- 266
- 5 Dasaratha V S, Richard C S, Eric B B. Process modeling using stacked neural networks. A IChE Journal, 1996, 42(9): 2529- 2539
- 6 Wolpert D H. Stacked generalization. Neural Networks, 1992, 5(2): 241- 259
- 7 Cho S B, Kim J H. Combining multiple neural networks by fuzzy integral for recognition. IEEE Trans on Syst, Man and Cybern, 1995, 25(2): 380- 384
- 8 Bates J M, Granger C W J. The combination of forecasts. Oper Res Q, 1969, 20(4): 451
- 9 Hunt K J, Sbarbaro D, Zbikowski R *et al*. Neural networks for control systems—A survey. Automatica, 1992, 28(6): 1083- 1112
- 10 罗荣富, 邵惠鹤, 王林, 等. 丙烯精馏塔的非线性推断控制系统. 化工自动化及仪表, 1992, 19(5): 5- 9
- 11 梁旻, 秦永胜, 徐用懋. 大型多元精馏塔控制系统的仿真. 化工自动化与仪表, 1996, 23(2): 23- 28

作 者 简 介

熊智华 男, 1971 年生。1996 年在东北大学自动控制系获硕士学位, 现为清华大学自动化系博士研究生。研究方向为复杂工业过程的建模与控制, 神经网络, 智能控制等。

王 雄 男, 1945 年生。1980 年在清华大学自动化系控制理论及应用专业获硕士学位, 现为清华大学自动化系教授。研究方向为复杂系统的控制和连续工业 C M S 等。

徐用懋 男, 1935 年生。1958 年毕业于清华大学动力系, 现为清华大学自动化系教授, 博士生导师。研究方向为生产过程建模与优化, 先进控制和连续工业 C M S 等。