

混合智能在递归神经网络软测量中的应用

王 杰¹, 古奋飞²

(1. 马鞍山师范高等专科学校, 安徽 马鞍山 243041; 2. 安徽新华学院信息工程学院, 合肥 230088)

摘 要:针对机理不清且具有复杂非线性关系的过程对象,提出一种基于混合智能的递归神经网络的数据驱动软测量建模方法。利用主元分析处理输入变量之间的共线性问题,设计了 Hammerstein 模型与递归神经网络的混合结构,用于描述过程对象的非线性和动态关系。为了使模型参数估计达到全局最优,提出了基于收缩粒子群优化算法的参数学习算法,不仅能够实现模型参数的更新,而且能够保证建模误差最小化。通过某氧化铝氯酸钠溶液组分浓度软测量问题中的实际应用表明:混合智能方法有效。

关键词:混合智能; 数据驱动; 递归神经网络; 粒子群优化算法

中图分类号:TP18

文献标志码:A

引 言

在实际应用中,难免会出现一些应用计算机技术难以直接测量的变量,对于其中的重要变量,常采用估算或推断的方式求解。软测量就是通过寻找一些相对容易测量的变量并建立它们与重要变量的数学关系,以此来对这些重要变量进行估算或推断的测量方法,软测量很好地实现了软件对硬件的替代。针对机理不清的过程对象,一般采用基于数据驱动(DD)的软测量建模方法。对于主变量和辅助变量之间的数学关系,基于 DD 的建模方法是通过取自输入和输出数据的有用信息来进行构建的,由于无需太多过程知识,所以成为一种常用的软测量建模方法。当前研究的方法主要有自回归模型^[1]、人工神经网络模型^[2]、模糊模型^[3]、贝叶斯网络模型^[4]、主成分分析(PCA)^[5]、偏最小二乘(PLS)模型^[6]、支持向量机模型^[7]等。PCA 方法在软测量建模中应用普遍,能够处理变量间共线性问题,然而它只能提取数据中的线性信息。为了处理非线性问题,一些学者提出了混合 PCA 与神经网络、核函数、二次函数、样条函

数、Chebyshev 多项式、小波函数的混合建模方法^[8-13]。上述方法采用的原理分为两类,一类是外部对样本(输入矩阵)进行变换,另一类是保留 PCA 方法的线性外部模型,内部采用非线性模型进行拟合。

然而,当前软测量建模方法存在无法反映对象动态特性,对于复杂非线性关系拟合精度低等问题。为了解决这一问题,提出混合 PCA 算法、Hammerstein 模型以及递归神经网络(RNN)的新模型结构。利用 PCA 算法可以降低输入数据维数、去除共线性的优点,并且结合 Hammerstein 模型对数据的某些动态特征具有较好的描述能力和 RNN 在非线性函数方面的良好逼近能力的优点,提出一种新的非线性动态 PCA 建模方法,即 HRN-NPCA 建模方法。并且,通过氧化铝厂实例对所提方法进行了验证。

1 HRN-NPCA 模型结构

HRN-NPCA 软测量模型结构如图 1 所示,在模型的外部,通过 PCA 算法不仅实现了最优成分在样本数据中的成对提取,从而达到多变量系统降维的目的,同时还

收稿日期:2015-02-13

基金项目:安徽省质量工程示范实验实训中心项目(20101233)

作者简介:王 杰(1980-),男,安徽蚌埠人,讲师,硕士,主要从事软件开发与大数据应用方面的研究,(E-mail)842669898@qq.com;

古奋飞(1982-),男,安徽芜湖人,讲师,硕士生,主要从事无线传感器网络与数据挖掘方面的研究,(E-mail)gff123@136.com

消除了它们的共线性关系。在模型的内部,通过类似 Hammerstein 模型结构的 HRNN 网络对提取的每对成分进行建模,利用递归神经网络模型的特点,结合 PCA 方法,共同描述系统的非线性及动态特性。

1.1 外部 PCA 模型

作为一种统计分析方法,PCA 可以通过对多元事物

进行分析,提取事物的主要矛盾及其本质成分,实现事物的降维,从而简化问题的复杂性。本文利用 PCA 技术对原始数据进行信息抽取,在保证信息损失最小的前提下,消除冗余信息、去除共线性,以精简模型结构,提高模型运算速度。

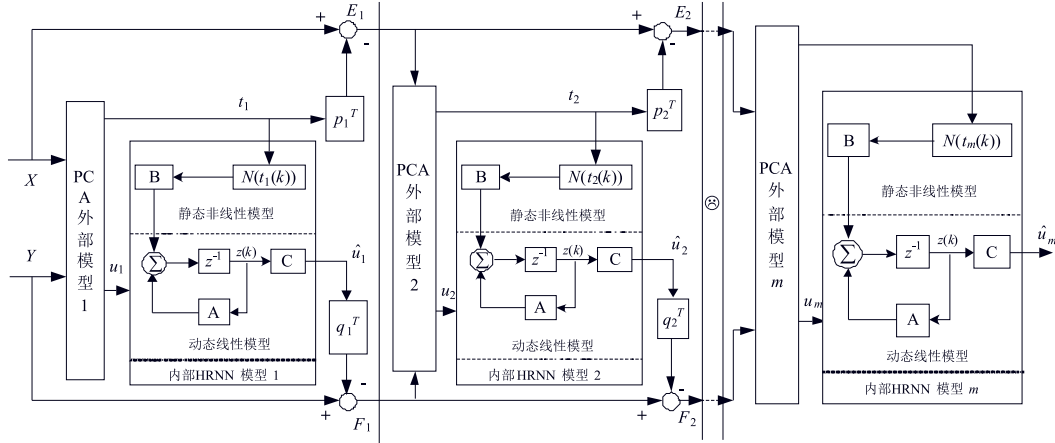


图 1 基于 HRNNPCA 方法的建模结构

1.2 内部 HRNN 模型

内部的 HRNN 模型是将 Hammerstein 模型与递归神经网络相结合的一种形式(图 2)。分成静态非线性和动态线性两部分。

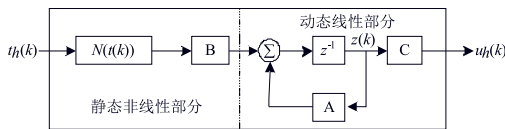


图 2 内部 HRNN 模型结构图

要获得静态非线性模型的输出空间,需要两步:首先利用非线性变换实现模型输入空间向状态空间的映射,然后使用一个线性动态变换实现状态空间向输出空间的映射。状态空间方程:

$$z(k + 1) = Az(k) + BN(t(k)) \quad (1)$$

$$u(k) = Cz(k) \quad (2)$$

其中: $A \in R^{J \times J}$, $B \in R^{J \times J}$, $C \in R^{1 \times J}$, $N \in R^J$ 。以第 h 个内部模型为例, $t = [t_h]$, $u = [u_h]$, $h = 1, 2, \dots, m$, 表示第 h 个内部模型的输入和输出变量。

Hammerstein 模型是一个由线性系统(NS)与一个非线性增益(NLG)相互串接而成的模型^[14],该模型通过 NS 对对象的动态特性进行描述,通过 NLG 对线性模型进行校正。由于 Hammerstein 模型非线性部分无需历史输入/输出信息,较易辨识、计算量少、能较好地反映过程特性,因此应用广泛。

递归神经网络采用递归预报误差算法训练神经网络,具有收敛速度快、收敛精度高的特点^[15]。由于其反

馈特征,使得递归神经网络模型能获取系统的动态响应特性,特别适用于非线性动态建模。

HRNN 网络继承了 Hammerstein 模型和神经网络的优点。以第 h 对主成分建立的内部 HRNN 模型为例,其网络结构如图 3 所示。

其中 $t_h(k)$ 和 $u_h(k)$ 分别表示第 h ($h = 1, 2, \dots, m$) 个内部模型的输入和输出变量。输出量 $u_h(k)$ 和状态变量 $z_h(k)$ 通过每层节点计算获得,公式如下:

$$u_h(k) = c_h z_h(k) = \sum_{j=1}^J c_{hj} z_{hj}(k) \quad (3)$$

$$z_{hj}(k) = \sum_{i=1}^J (a_{n_{ji}} z_{hi}(k-1)) + b_{n_{ji}} n_{ij}(k-1) \quad (4)$$

$$n_{ij}(k) = f(g_{ij}(k)) = \frac{\exp(g_{ij}(k)) - \exp(-g_{ij}(k))}{\exp(g_{ij}(k)) + \exp(-g_{ij}(k))} \quad (5)$$

$$g_{ij}(k) = w_{ij} t_h(k) + d_{ij} \quad (6)$$

其中: w_{ij} 用于表示隐含层第 j 个神经元与输入层之间的权值, d_{ij} 代表隐含层第 j 个节点对应的阈值。

2 HRNNPCA 模型算法

在 HRNNPCA 模型中,外部采用标准 PCA 算法获取特征向量,而对于内部 HRNN 模型参数,本文设计了基于智能算法的稳定学习参数估计算法。以第 h 个内部模型为例,模型输出误差为:

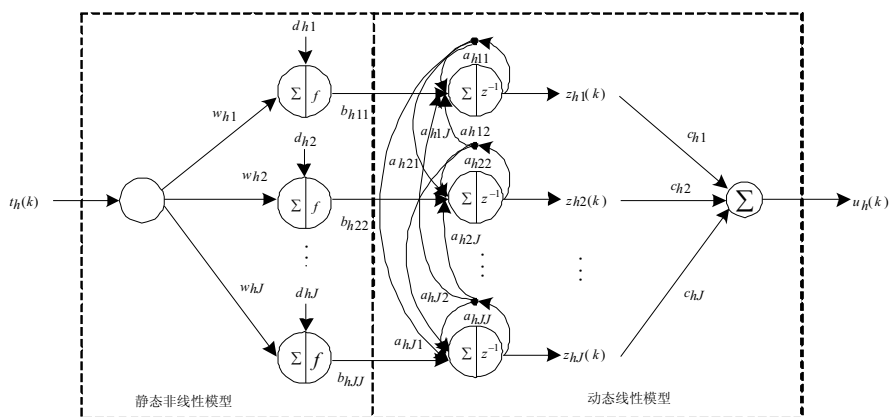


图3 内部 HRNN 模型拓扑结构

$$H(w, k) = \frac{1}{2}(u_h(k) - \hat{u}_h(k))^2 = \frac{1}{2}e_h(k)^2 \quad (7)$$

其中, $e_h(k) = u_h(k) - \hat{u}_h(k)$ 。

这里参数估计的准则为:

$$J = \min: H \quad (8)$$

本文采用收缩 PSO 算法实现,其步骤为:

第一步:初始化。

以 Pop 表示粒子种群规模,第 i 个粒子的初始位置通过 p_i 表示,而 $\{p_i | i = 1, \dots, Pop\}$ 表示 Pop 个粒子的初始位置组成的集合,集合中的所有点随机生成于决策空间。将所有粒子的初始速度设定为 0,即 $v_i = 0, i = 1, \dots, Pop$ 。以粒子的自身位置对所有粒子的个体向导进行初始化,第 i 个粒子的个体向导 l_i 初始化为自身位置即 $l_i = p_i$,以随机数对全局向导 g 的初始位置进行初始化。

第二步:对所有粒子的目标值进行评估。

$H_i, i = 1, \dots, Pop$, 表示将第 i 个粒子的当前位置代入式(8)中所得值,可以用于表示该粒子的当前位置的目标值,该目标值的取值越小越好。

第三步:对向导进行调整。

假设 H_g 是全局向导 g 的目标值,而 H_l_i 表示第 i 个粒子对应的个体向导 l_i 的目标值,对 PSO 种群的个体向导与全局向导进行调整的公式为:

$$l_i \leftarrow \begin{cases} l_i & \text{if } H_i \geq H_l_i \\ p_i & \text{if } H_i < H_l_i \end{cases} \quad (9)$$

$$g \leftarrow \begin{cases} g, & \text{if } H_g \leq \min_{i=1}^{Pop} \{H_i\} \\ \min_{i=1}^{Pop} \{H_i\}, & \text{if } H_g > \min_{i=1}^{Pop} \{H_i\} \end{cases} \quad (10)$$

第四步:对粒子的位置和速度进行更新。

更新所依据的公式为:

$$\begin{cases} v_i \leftarrow \chi[v_i + c_1c_3(l_i - p_i) + c_2c_4(g - p_i)] \\ p_i \leftarrow v_i + p_i \end{cases} \quad (11)$$

式中, c_1 与 c_2 分别为加速因子参数,且 $c_1 + c_2 > 4, \chi = 2 / |2 - (c_1 + c_2) - \sqrt{(c_1 + c_2)^2 - 4(c_1 + c_2)}|$ 。 c_3 与 c_4 表示随机分布于区间 (01) 内的两个数。

第五步:对是否满足算法的搜索停止条件进行判断。

判断的依据:如果迭代次数 G 没有达到设定的最大搜索次数,返回第二步继续进行寻优;否则,停止搜索,并将搜索到的参数最优值输出。

3 实验研究

3.1 实验对象与数据

将提出的基于 HRNNPCA 的非线性动态建模方法用于建立铝酸钠溶液组分浓度软测量模型。

经过数据预处理的 380 组数据作为组分浓度软测量模型的离线训练数据,用来建立铝酸钠溶液组分浓度的软测量模型,预留的 150 组样本数据用来测试软测量方法的有效性,部分测试数据见表 1。

表 1 测试数据

变量	$T_1 / ^\circ\text{C}$	$d_1 / (\text{ms/cm})$	$T_2 / ^\circ\text{C}$	$d_2 / (\text{ms/cm})$	$T_3 / ^\circ\text{C}$	$d_3 / (\text{ms/cm})$	$c_K / (\text{g/l})$	$c_A / (\text{g/l})$	$c_C / (\text{g/l})$
1 组	86.9	572.58	72.6	460.63	79	503.05	205	99.7	29.6
2 组	86.5	577.73	72.4	463.83	78	507.66	207	99.7	31.2
3 组	86.4	568.91	73.5	470.47	79	504.06	205	98.0	29.6
4 组	91.0	606.63	76.0	468.56	83	537.81	203	100.0	30.8
...
149 组	92.9	660.70	74.2	486.33	79	521.41	218	104.0	29.8
150 组	96.1	701.72	72.5	463.20	80	552.19	217	103.0	29.0

3.2 模型参数

选择铝酸钠溶液的三组温度和电导率 T_1 、 d_1 、 T_2 、 d_2 、 T_3 、 d_3 作为组分浓度模型的输入变量,即三种组分浓度的软测量模型均有 6 个输入变量、1 个输出变量。首

先在外利用 PCA 算法提取特征向量,得到消除共线性后的样本数据集,计算模型对其的预测误差。HRN-NPCA 模型累计方差百分比计算结果见表 2。

表 2 HRNNPCA 模型累计方差百分比

LV#	c_K				c_A				c_C			
	X - Block		Y - Block		X - Block		Y - Block		X - Block		Y - Block	
	This LV	Total	This LV	Total	This LV	Total	This LV	Total	This LV	Total	This LV	Total
1	95.83	95.67	90.19	90.10	95.92	95.88	93.84	93.86	95.92	95.84	89.02	87.14
2	1.73	97.61	0.02	90.13	0.66	96.54	0.00	93.67	2.61	98.61	0.00	87.14

内部神经网络结构设计一般依赖于人的经验,也有一些经验取值法,比如数据样本数为 N , 隐含层节点个数 H 满足 $N \leq 2^H$ 。一般情况下,隐含层节点数越多,精度越高,但隐含层节点个数过多,模型过于复杂,就会出现参数“过拟合”现象。由于内部模型的递归神经网络是单入单出的,经过试验,这里选择隐层节点个数为 7 个,即内部 HRNN 网络结构为 1-7-7-1。内部递归神经网络第一个隐含层权值矩阵 W , 隐含层阈值矩阵 W_D , 第二个隐含层权值矩阵 W_B , 第二个隐含层另一权值矩阵 W_A , 输出层权值向量 W_C 。经过收缩粒子群算法优化,神经网络参数选为 $\eta_h = 1.2$ 。

3.3 模型结果与验证

按照本文提出的 HRNNPCA 建模方法,苛性碱浓度模型的训练和测试结果如图 4 所示,误差自相关函数如图 5 所示。从测试结果曲线和误差自相关函数曲线可以看出,氧化铝浓度计算值曲线趋势正确,且精度较高。

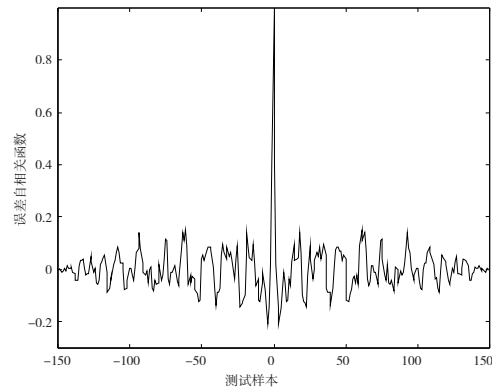


图 5 c_K 测试误差自相关函数

部分无需历史输入、输出信息,较易辨识、计算量少、能较好地反映过程特性。将所提方法应用到某氧化铝氯酸钠溶液组分浓度软测量问题中,应用结果验证了所提方法的有效性。

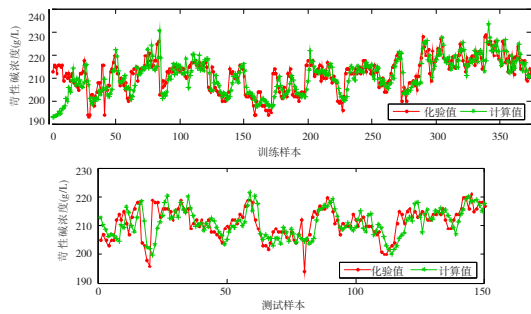


图 4 基于 HRNNPCA 的苛性碱浓度训练和测试结果

4 结束语

本文在前人研究的基础上提出了一种基于混合的 RNN 的 DD 的建模方法。该方法针对氧化铝氯酸钠溶液组分浓度软测量这一类具有动态、非线性特性的工业对象建模问题,将对象分解成一个线性动态系统和一个非线性即时系统,从而建立一个由 NS 与一个 NLG 相互串接而成的模型,该模型通过 NS 对对象的动态特性进行描述,通过 NLG 对线性模型进行校正。该模型非线性

参考文献:

- [1] Petr K, Bogdan G, Sibylle S. Data - driven soft sensors in the process industry[J]. Computers and Chemical Engineering, 2009, 33(4):795-814.
- [2] Chen C, Yan X. Optimization of a multilayer neural network by using minimal redundancy maximal relevance-partial mutual information clustering with least square regression[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014(99):1.
- [3] Zhou Ping, Lu Shaowen, Chai Tianyou. Data-Driven soft-sensor modeling for product quality estimation using case-based reasoning and fuzzy-similarity rough sets[J]. IEEE Transactions on Automation Science and Engineering, 2014, 11(4):992-1003.
- [4] Petr K, Ratko G, Bogdan G. Review of adaptation mechanisms for data-driven soft sensors[J]. Computers and Chemical Engineering, 2011, 35(1):1-24.

- [5] Wang Jianguo, Xiao Qianping, Wang Juanuan, et al. Soft sensor approach for optimal operation of cooling tower for energy conservation[C]//Proceeding of The 26th Chinese Control and Decision Conference (2014 CCDC), Changsha, China, May 31-Jun 2, 2014: 3612-3615.
- [6] Lin B, Recke B, Knudsen J, et al. A systematic approach for soft sensor development[J]. Computers and Chemical Engineering, 2007, 31(5): 419-425.
- [7] Shang Chao, Gao Xinqing, Yang Fan, et al. Novel Bayesian framework for dynamic soft sensor based on support vector machine with finite impulse response[J]. IEEE Transactions on Control Systems Technology, 2014, 22(4): 1550-1557.
- [8] 黄勤, 常伟, 刘益良, 等. 基于 PCA 的 BP 神经网络分类器[J]. 重庆工学院学报: 自然科学版, 2009(7): 354-356.
- [9] Muller K, Mika S, Ratsch G, et al. An introduction to kernel-based learning algorithms[J]. IEEE Transactions on Neural Networks, 2001, 12(2): 181-201.
- [10] Dunia R, Qin J, Edgar T F, et al. Identification of faulty sensors using principal component analysis[J]. AIChE Journal, 1996, 42(10): 2797-2812.
- [11] Goethals I, Pelckmans K, Suykens J A, et al. Subspace identification of Hammerstein systems using least squares support vector machines[J]. IEEE transactions on Automatic Control, 2005, 50(10): 1509-1519.
- [12] 巩耀晓, 杨文考, 范五东. 对称分数 B 样条小波的 PCA 图像融合[J]. 计算机工程与应用, 2012, 4(3): 213-215.
- [13] Stoica P. On the convergence of an iterative algorithm used for Hammerstein system identification[J]. IEEE Trans on Automatic Control, 1981, 26(4): 967-969.
- [14] Bai E. An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems[J]. Automatica, 1998, 34(3): 333-338.
- [15] Wang J S, Chen Y P. A fully automated recurrent neural network for unknown dynamic system identification and control[J]. IEEE Transactions on Circuits and Systems-I, 2006, 56(6): 1363-1372.

Application of Hybrid Intelligent in Recurrent Neural Network Soft Measurement

WANG Jie¹, GU Fenfei²

(1. Maanshan Teacher's College, Maanshan 243041, China; 2. College of Information Engineering, Anhui Xinhua University, Hefei 230088, China)

Abstract: For the process objects with unclear mechanism and complex nonlinear relationship, a data-driven soft sensor modeling method of recurrent neural network is proposed based on hybrid intelligent. Through the principal component analysis, the co-linear problem among input variables is processed, then the hybrid structure containing Hammerstein model and recurrent neural network is designed, which is used to describe the nonlinear and dynamic relationship of process object. In order to get the global optimum of model parameters estimation, the parameters learning algorithm based on the contraction PSO is proposed, it can not only update the model parameters, but also ensure the modeling errors be minimized. The proposed method is applied to the concentration soft measurement of aluminum sodium chlorate component, and the application results show the effectivity of the hybrid intelligent method.

Key words: hybrid intelligent; data-driven; recurrent neural network; Particle Swarm Optimization