

改进 RELM 在多变量解耦控制中的应用

丁金林¹, 王 峰^{1,2}, 孙 洪¹, 刘国海²

(1 苏州市职业大学 电子工程系, 江苏 苏州 215104; 2 江苏大学 电气信息工程学院, 江苏 镇江 212013)

摘 要: 针对神经网络逆系统方法实现复杂非线性系统解耦存在训练时间长、实时控制较差的缺陷, 提出一种改进的 RELM(正则极限学习机)训练算法, 根据输出权值的特点, 采用不带平方根的乔累斯基分解, 提高了计算效率, 减少了训练时间, 具有较高的学习精度及更好的泛化能力; 进一步将此神经网络应用到 3 输入 3 输出多变量离散系统的解耦控制, 仿真实验结果表明, 所提出的方法具有较快的实时控制速度, 具有较高的实用价值。

关键词: 神经网络; 非线性; 正则极限学习机; 乔累斯基分解; 逆系统; 解耦控制

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2012)11-0070-04

Multivariate Decoupling Control Based on Improved Regularized Extreme Learning Machine

DING Jin-lin¹, WANG Feng^{1,2}, SUN Hong¹, LIU Guo-hai²

(1 Suzhou Vocational University, Suzhou 215104, China;

2 School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China)

Abstract: An improved RELM learning algorithm is prospected to solve the long-time training and the poor real-time control in complex nonlinear decoupling control based on neural network inverse method. The RELM with higher learning accuracy and better generalization ability can improve computational efficiency and reduce training time by use of cholesky factorization without square root based on the feature of output weights. And the method is further applied in decoupling control of multivariate discrete system with 3-input and 3-output. Simulation results demonstrate that the proposed method has faster real-time control speed, and has very high practical value.

Key words: neural network; nonlinear; rELM; cholesky factorization; inverse system; decoupling control

1 引言

对于多输入多输出非线性离散系统的解耦, 逆方法是一种有效的控制方法^[1]. 在非线性解耦控制中, 逆模型辨识的速度、精度及泛化能力是影响系统解耦效果的关键. 不同的神经网络其性能差别比较大, 传统的前向神经网络(如 BP)存在网络训练时间长, 过拟合、不能保证全局最优性等问题^[2-3]. 支持向量机回归算法^[4](SVM)是一种基于统计学习理论的学习机, 避免了神经网络中可能出现的“过拟合”现象, 但其适合于小样本学习, 当样本数目很大时,

其学习时间过长, 另外对于支持向量机不确定性参数(正则化参数、核参数、不灵敏参数)的优化选择和确定问题, 没有很好的解决方法.

基于极限学习机(ELM)的单隐层前向神经网络(SLFNS)具有学习速度快等优点^[5-8], 可用于非线性系统控制. 但由于 ELM 是基于结构风险最小化原理, 可能会导致过度拟合问题. 此外 ELM 不考虑误差的权重, 当数据集存在离群点时, 它的性能将会受到严重影响^[8]. 为了克服这些缺点, 结合结构风险最小化理论及加权最小二乘法对 ELM 算法进行改进, 提出正则极限学习机(RELM), 泛化性能得到

收稿日期: 2012-05-14; 修回日期: 2012-06-15

基金项目: 国家自然科学基金项目(51077066); 江苏省自然科学基金项目(BK2011319); 苏州市职业大学创新基金资助项目(2011SZDCC01)

进一步的提高。

正则极限学习机,能实现复杂非线性系统的实时控制,但其通过求逆的方法计算输出矩阵^[9-10],其训练精度、泛化能力受较大影响,学习时间相对较长。本文通过分析 RELM 学习算法的特点,提出了基于改进的 RELM 训练算法。

2 正则极限学习机(RELM)

根据统计学理论可知,实际风险包括经验风险和结构风险两种成份^[5]。RELM 同时考虑这两种风险因素,并通过参数 γ 调节两种风险的比例,RELM 的数学模型可表示为

$$\begin{aligned} \arg\min_{\beta} E(W) &= \arg\min_{\beta} \left[\frac{1}{2} \gamma \|\epsilon\|^2 + \frac{1}{2} \|\beta\|^2 \right] \\ \text{s. t. } \sum_{i=1}^M \beta_i g(a_i \cdot X_j + b_i) - t_i &= \epsilon_j \\ j &= 1, 2, \dots, N \end{aligned} \quad (1)$$

其中,不同样本的误差加权平方和 $\|\epsilon\|^2$ 表示经验风险, $\|\beta\|^2$ 代表结构风险,而 γ 则是两种风险的比例参数,通过交叉验证的方式确定 γ ,由此获得两种风险的最佳折中点。

构造拉格朗日方程

$$\xi(\beta, \epsilon, \alpha) = \frac{1}{2} \gamma \|\epsilon\|^2 + \frac{1}{2} \|\beta\|^2 - \alpha(H\beta - T - \epsilon)$$

其中 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_N]$; $\alpha_j \in R^m$, ($j=1, 2, \dots, N$) 代表拉格朗日乘子。拉格朗日方程分别对 β, ϵ, α 计算偏导,求解得:

$$\begin{cases} \beta^T = \alpha H \\ \gamma \epsilon^T + \alpha = 0 \\ H\beta - T - \epsilon = 0 \end{cases} \quad (2)$$

故 $\beta = (\gamma^{-1}I + H^T H)^T H^T T$, I 为单位矩阵,利用 β 可得到基于 RELM 的非线性解耦控制模型 $t = \sum_{i=1}^M \beta_i g(a_i \cdot x + b_i)$, 式中 x 为模型输入, t 为模型输出。

3 改进的 RELM 算法

在进行 RELM 训练时,主要求解输出权值 β 。当样本数量比较大,计算量大,影响训练速度时,根据 RELM 输出权值矩阵的特点,提出采用不带平方根的乔累斯基分解,可提高训练速度,减少训练时间。

由式(2)得

$$(\gamma^{-1}I + H^T H)\beta = H^T T \quad (3)$$

令 $A = \gamma^{-1}I + H^T H$, $b = H^T T$, 则式(3)变为

$$A\beta = b \quad (4)$$

求解线性方程组(4)得到 β 。

由于 $A^T = (\gamma^{-1}I + H^T H)^T = \gamma^{-1}I + H^T H = A$, 故 A 为对称矩阵。

对于任意向量 $x = (x_1, x_2, \dots)^T$, 其二次型为

$$\begin{aligned} x^T A x &= x^T (\gamma^{-1}I + H^T H) x \\ &= \gamma^{-1} x^T x + (Hx)^T Hx \end{aligned} \quad (5)$$

显然当 $x \neq 0$ 时, $\gamma^{-1} x^T x > 0$, $(Hx)^T Hx > 0$, 故 $x^T A x > 0$, 则 A 为正定矩阵。综上所述,可知 A 为对称正定矩阵,可进行不带平方根乔累斯基分解:

$$A = LDL^T \quad (6)$$

其中 L 为具有正的对角线元素的下三角矩阵, D 为对角阵。

利用不带平方根乔累斯基分解求 RELM 输出权值 β 的步骤如下:

① 利用不带平方根乔累斯基分解计算 L, D , L 矩阵中不为零的元素可利用矩阵 A 中元素 a_{ij} 根据

$$l_{ij} = a_{ij} - \sum_{k=1}^{j-1} \frac{l_{ik} l_{jk}}{l_{kk}} \text{ 计算, 对角矩阵 } D \text{ 中元素}$$

$$d_{ii} = \frac{1}{l_{ii}}, \text{ 其中 } i=1 \dots m, \quad j=1 \dots i;$$

② 将式(6)代入式(4)得: $LDL^T \beta = b$, 两边乘以 $D^{-1}L^{-1}$ 得:

$$L^T \beta = D^{-1}L^{-1}b \quad (7)$$

令 $F = D^{-1}L^{-1}b$, 则 $LDF = b$, 通过比较两端元素, 矩阵 F 中的元素 f_{ij} 的计算公式

$$f_{ij} = \begin{cases} b_{ij} & i = 1 \\ b_{ij} - \sum_{n=1}^{i-1} \frac{l_{in} f_{nj}}{l_{nn}} & i > 1 \end{cases} \quad (8)$$

③ 由 $L^T \beta = F$, 得到输出矩阵 β 中的元素 β_{ij} :

$$\beta_{ij} = \begin{cases} \frac{f_{ij}}{l_{ii}} & i = m \\ \frac{(f_{ij} - \sum_{n=1}^{m-i} l_{i+n,i} \beta_{i+n,j})}{l_{ii}} & i < m \end{cases} \quad (9)$$

4 基于改进的 RELM 多变量系统解耦控制实例

4.1 多变量系统实例分析

复杂 3 输入 3 输出非线性时延离散系统, 系统的传递函数为

$$\begin{cases} y_1(k+1)=0.4y_1(k)+u_1(k)/[1+u_1(k)^2]+ \\ \quad 0.2u_1(k)^3+0.5u_2(k)+0.3y_2(k) \\ y_2(k+1)=0.2y_2(k)+u_2(k)/[1+u_2(k)^2]+ \\ \quad 0.4u_2(k)^3+0.2u_1(k)+0.3y_3(k) \\ y_3(k+1)=0.3y_3(k)+u_3(k)/[1+u_3(k)^2]+ \\ \quad 0.4u_3(k)^3+0.4u_2(k)+0.3y_1(k) \end{cases} \quad (10)$$

其中 u_1 、 u_2 、 u_3 为系统的三个输入量, y_1 、 y_2 、 y_3 为控制量, 对原系统同时输入 3 个相位相差 90° , 幅度皆为 0.25 的方波信号, 如图 1 中虚线 u_1 、 u_2 、 u_3 所示, 其输出响应曲线 y_1 、 y_2 、 y_3 分别如图 1 中实线所示, 可看出该系统是强耦合非线性。

$$\frac{\partial F}{\partial [u_1(k), u_2(k), u_3(k)]} = \begin{bmatrix} \frac{1+1.6u_1^2(k)+u_1^4(k)}{-(1+u_1^2(k))^2} & 0.5 & 0 \\ 0.2 & \frac{1+2.2u_2^2(k)+u_2^4(k)}{-(1+u_2^2(k))^2} & 0 \\ 0 & 0.4 & \frac{1+2.2u_3^2(k)+u_3^4(k)}{-(1+u_3^2(k))^2} \end{bmatrix} \quad (11)$$

显然: $\det \frac{\partial F}{\partial [u_1(k), u_2(k), u_3(k)]} \neq 0$, 式(10)所示的离散系统可逆, 构造神经网络逆, 且 $\alpha_1=1, m_1=0, \alpha_2=1, m_2=0, \alpha_3=1, m_3=0$ 。

4.2 基于改进的 RELM 神经网络控制器构造及仿真分析

构造基于不带平方根乔累斯基分解的 RELM 神经网络 α 阶逆控制器, 输入层节点数为 6, 输出层节点数为 3, 隐含层节点数为 22, 对原系统输入同时施加 3 路互不相关白噪声, 得到输入输出数据 10000 组, 将其重新组合神经网络训练样本数据集, 即 $\{y_1(l+1), y_1(l), y_2(l+1), y_2(l), y_3(l+1), y_3(l)\}$ 和 $\{u_1(l), u_2(l), u_3(l)\}$, 前 8000 组用于神经网络的训练样本, 后 2000 组样本用于测试。随机给定神经网络输入权值 a_i 、偏差 $b_i (i=1, 2, \dots, \bar{N})$, 激励函数选“Sigmoid”函数, 采用改进的 RELM 网络进行训练与测试的仿真结果如表 1 所示。为了进一步证明该方法的有效性, 同时采用 ELM、RELM、BP、SVM 等方法进行仿真, 结果见表 1。

对表 1 中数据进行分析可知:

① ELM、RELM、改进 RELM 的训练时间明显小于传统的 BP 和 SVM, 速度较快, 其中改进 RELM 训练所需时间最短;

② ELM、RELM、改进 RELM 三者训练精度相

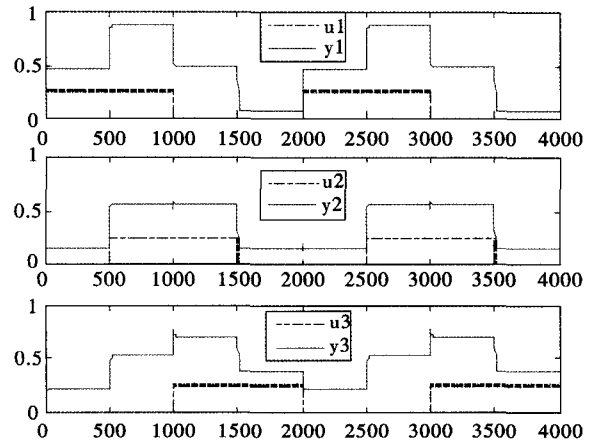


图 1 原离散系统方波响应曲线

定义 $F(\cdot)$ 为式(10)中等号右边移到左边构成的函数, 对 $U(k)$ 求偏导, 有:

表 1 不同神经网络性能比较

神经网络类型	学习时间 (s)	测试时间 (s)	训练精度 (RMSE)	测试精度 (RMSE)	层数/向量数
ELM	0.1063	0.0241	0.0194	0.0457	22
	0.2838	0.0509	0.0182	0.0415	50
RELM	0.8934	0.0822	0.0136	0.0158	100
	1.8675	0.1472	0.0062	0.0138	200
	0.0681	0.0250	0.0194	0.0461	22
	0.1297	0.0488	0.0184	0.0401	50
	0.2719	0.0759	0.0172	0.0217	100
改进	0.6571	0.1397	0.0150	0.0267	200
	0.0676	0.0275	0.0201	0.0400	22
	0.1263	0.0444	0.0153	0.0323	50
RELM	0.2687	0.0725	0.0150	0.0258	100
BP	0.6447	0.1363	0.0129	0.0181	200
	65.4823	0.8584	0.0565	0.0712	22
SVM	112.563	5.3625	0.0465	0.0672	2000

近, 但泛化能力差别较大, 由于 ELM 其只考虑经验风险, 未考虑结构风险, 泛化能力比较差, 改进 RELM 的泛化能力最强;

③ 由于改进 RELM 输出权值采用不带平方根

乔累斯基分解,只使用简单的四则运算,避免了 RELM 和 ELM 中的求逆,训练样本时间更短,特别是大样本复杂非线性系统更加明显;

④ 故选用改进 RELM 神经网络对式(10)系统进行解耦。

将训练好的基于改进 RELM 的神经网络逆系统串联在原被控对象之前,构成伪线性复合控制系统,可实现控制量 y_1 、 y_2 、 y_3 之间的解耦控制,其构成框图如图 2 所示。

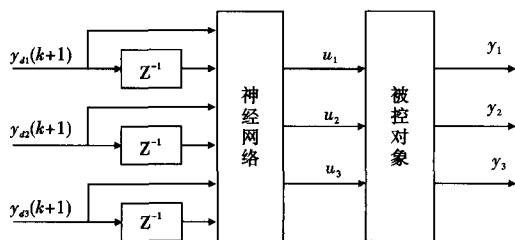


图 2 基于改进 RELM 的神经网络解耦控制系统

为检验伪线性复合系统解耦效果,给伪线性复合系统的 3 个输入端 y_{d1} 、 y_{d2} 、 y_{d3} 分别同时输入如图 3 虚线所示相位幅度不同方波信号组,其中 y_{d1} 为幅度 0.7 的方波, y_{d2} 为幅度 0.4 的方波, y_{d3} 为幅度 0.6 的方波,相位依次相差 90° ,系统的响应曲线分别如图 3 实线 y_1 、 y_2 、 y_3 所示。显然复合系统已经基本解耦为 3 个独立的 SISO 伪线性离散系统。

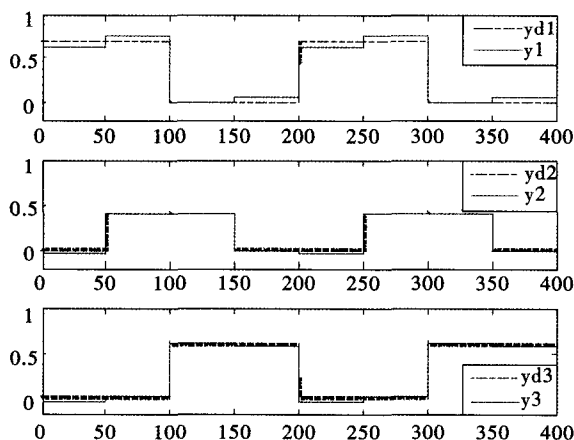


图 3 基于改进 RELM 的解耦图

5 结束语

本文提出了一种新的基于改进 RELM 训练算法,其根据输出权值特点,采用不带平方根乔累斯基分解,提高了网络训练的速度,并且有较好的学习精度和更强的泛化能力,适用于复杂非线性解耦控制。通过一个 3 输入 3 输出的复杂非线性系统实例分析,使用基于改进的 RELM 神经网络逼近离散非线

性系统的逆系统,并将此逆系统串联在原系统之前,形成伪线性复合系统,将原系统转化为 3 个单输入单输出的线性系统,实现非线性系统的线性化解耦。此控制方法可推广到其他复杂非线性系统控制,具有很好的应用前景。

参考文献:

- [1] Dai Xianzhong, Liu Guohai, Zhang Xinghua. Neural network inverse control of variable frequency speed-regulating system in v/f mode [J]. Proceedings of the CSEE, 2005, 25(7): 109-114.
- [2] Huang D S, Ji-Xiang Du. A constructive hybrid structure optimization methodology for radial basis probabilistic neural networks[J]. IEEE Transactions on Neural Networks, 2008, 19(12): 2099-2115.
- [3] Huang G B, Zhu Q Y, Siew C K. Real-time learning capability of neural networks[J]. IEEE Transactions on Neural Networks, 2006(17): 863-878.
- [4] Liu Guohai, Dong Beibei, Teng Chenglong, et al. Model reference adaptive control of PMSM based on support vector machines generalized inverse[J]. Journal of Southeast University: Natural Science Edition, 2010(40): 13-18.
- [5] Fei Han, Huang D S. Improved extreme learning machine for function approximation by encoding a priori information[J]. Neurocomputing, 2006 (69): 2369-2373.
- [6] 李彬,李贻斌. 基于 ELM 学习算法的混沌时间序列预测[J]. 天津大学学报, 2011, 44(8): 701-704.
- [7] 韩敏,李德才. 基于替代函数及贝叶斯框架的 1 范数 ELM 算法[J]. 自动化学报, 2011, 37(11): 1344-1350.
- [8] 邓万宇,郑庆华,陈琳,等. 神经网络极速学习方法研究[J]. 计算机学报, 2010, 33(2): 279-287.
- [9] Feng G, Huang G B, Lin Q, et al. Error minimized extreme learning machine with growth of hidden nodes and incremental learning[J]. IEEE Transactions on Neural Networks, 2009, 20 (8): 1352-1357.
- [10] 张弦,王宏力. 基于 Cholesky 分解的增量式 RELM 及其在时间序列预测中的应用[J]. 物理学报, 2011, 60 (11): 1-6.

作者简介:



丁金林 女, (1978—), 硕士, 讲师. 研究方向为自动化控制。