第44卷 第5期 电 子 与 信 息 学 报 Vol. 44No. 5
2022年5月 Journal of Electronics & Information Technology May 2022

哈默斯坦非线性时变系统的加权学习辨识方法

仲国民* 俞其乐 陈 强
(浙江工业大学信息工程学院 杭州 310023)

摘 要：针对有限区间哈默斯坦(Hammerstein)非线性时变系统，该文提出一种加权迭代学习算法用以估计系统时变参数。首先将Hammerstein系统输入非线性部分进行多项式展开，采用迭代学习最小二乘算法辨识系统的时变参数。为了防止数据饱和，采用带遗忘因子的迭代学习最小二乘算法，进而引入权矩阵，采用加权迭代学习最小二乘算法改进系统跟踪误差，以提高辨识精度。该文分别给出3种算法的推导过程并进行仿真验证。结果表明，与迭代学习最小二乘算法和带遗忘因子迭代学习最小二乘算法相比，加权迭代学习最小二乘算法具有辨识精度高、跟踪误差小以及迭代次数少等优点。

关键词：加权迭代学习辨识；时变参数；哈默斯坦模型；最小二乘算法
中图分类号：TN911.7; TP181 **文献标识码**：A **文章编号**：1009-5896(2022)05-1610-07
DOI: 10.11999/JEIT210857

Weighted Learning Identification Method for Hammerstein Nonlinear Time-varying Systems

ZHONG Guomin YU Qile CHEN Qiang
(*College of Information Engineering*, *Zhejiang University of Technology*, *Hangzhou* 310023, *China*)

Abstract: For Hammerstein nonlinear time-varying systems running repeatedly on finite intervals, a weighted iterative learning algorithm is proposed to estimate the time-varying parameters involved in the system dynamics. The nonlinear input part of the Hammerstein system is tackled based on polynomial expansion, and the iterative learning least square algorithm is given for the time-varying parameter identification. In order to prevent data saturation, an iterative learning least squares algorithm with forgetting factor is proposed for reducing the system tracking error and improving the identification accuracy; A weighted iterative learning least squares algorithm is further presented by introducing the weight matrix. The derivations of the three algorithms are given in detail. The simulation results demonstrate the effectiveness of the proposed learning algorithms, and in comparison with iterative learning least squares algorithm, the modified one sreach high identification accuracy and need fewer iterations.

Key words: Weighted iterative learning identification; Time-varying parameters; Hammerstein Model; Least squares algorithm

1 引言

非线性特性广泛存在于实际工业生产过程中，为便于研究其变化规律，需要用数学模型来对所研究的物理现象或过程进行定量分析，因此非线性系统辨识问题受到越来越多的关注。非线性系统结构复杂，难以采用统一模型进行描述，在过去的几十年里，国内外学者主要致力于较普遍的块结构非线性系统辨识[1]。块结构非线性系统可以分为非线性静态部分和线性动态部分，分别用非线性函数和线性函数来表示其特性[2]。其中，维纳(Wiener)系统、哈默斯坦(Hammerstein)系统及其组合形式是应用最广泛的配置。哈默斯坦系统广泛应用于模拟连续搅拌釜式反应器、pH中和过程、蒸馏塔、液压自动发电量控制系统、多传感器系统等非线性过程或系统[3,4]。它由1个非线性静态函数串联和1个线性动态子系统构成，其中哈默斯坦受控自回归滑动平均模型(Controlled AutoregRessive Moving Average model, CARMA)系统的具体形式描述如图1所示。

针对哈默斯坦系统的辨识问题，已开展较多研究工作，如子空间辨识法[5–8]、过参数化法[9,10]、

收稿日期：2021-08-19；改回日期：2022-01-07；网络出版：2022-02-02
*通信作者： 仲国民 zgm@zjut.edu.cn
基金项目：国家自然科学基金(62073291, 62973274)
Foundation Items: The National Natural Science Foundation of China (62073291, 62973274)

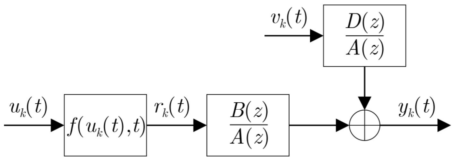


图 1 哈默斯坦系统CARMA模型

最小二乘法[11]、极大似然法[12]、非参数核回归估计[13,14]、分数阶法[15]、基于特殊输入信号的方法[16]和相关分析方法[17]等。文献[16]提出一种识别哈默斯坦非线性过程的方法，该方法利用一个特殊的测试信号，能够识别出哈默斯坦系统由1个非线性静态函数和1个线性动态子系统组成。针对哈默斯坦非线性带外部输入的自回归滑动平均模型(Autoreg-Ressive Moving Average model with eXogenous Input, ARMAX)，文献[18]提出了一种迭代梯度算法。上述方法均是假设静态非线性特性为子函数多项式的组合前提下，针对定常哈默斯坦系统开展的研究。现有带遗忘因子递推算法[19]、块脉冲函数法[20]等能够提高时变参数跟踪性能；迭代学习辨识算法[21,22]有效解决了时变参数估计问题。鉴于块结构时变非线性系统辨识方面的研究尚少，本文将迭代学习方法进一步运用于基于块结构的非线性系统辨识。

本文提出一种加权迭代学习辨识的方法，其动机是已有系统辨识算法加权修正的思想[23]。当动态系统在有限区间上重复运行时，沿重复轴来看，固定时刻对应的参数是相对固定的，可以采用迭代学习算法估计时变参数。但由于实际中存在非重复初始条件和外部干扰，一致重复运行并不是总能保证的。因此，为提高参数估计精度和系统跟踪效果，将时间轴上加权矩阵方法推广运用于沿重复轴上的修正，构建加权迭代学习算法。

本文考虑有限区间上重复运行的哈默斯坦非线性时变系统，借助哈默斯坦定常系统辨识方法，采用辅助模型方法[24]，推导出哈默斯坦非线性时变系统基于“重复轴”的迭代学习最小二乘算法。同时，为避免数据饱和，引入遗忘因子，推导出带遗忘因子迭代学习最小二乘算法。在此基础上，进一步改进准则函数，引入权矩阵，探索一种加权迭代学习最小二乘(Weighted Iterative Learning Least Squares, WILLS)的辨识方法，并将该算法运用于时变哈默斯坦模型的辨识研究，其优点是达到一定辨识精度的条件下，迭代次数少，输出误差小且稳定性好。

2　问题描述

考虑如式(1)所述有限区间上重复运行的单输入单输出离散时变哈默斯坦系统

$$A(q^{-1},t)y_k(t)=B(q^{-1},t)f(u_k(t),t)+D(q^{-1},t)v_k(t) \tag{1}$$

其中

$$A(q^{-1},t)=1+a_{1k}(t)q^{-1}+a_{2k}(t)q^{-1}+\cdots+a_{n_ak}(t)q^{-n_a} \tag{2}$$

$$B(q^{-1},t)=b_{1k}(t)q^{-1}+b_{2k}(t)q^{-1}+\cdots+b_{n_bk}(t)q^{-n_b} \tag{3}$$

$$D(q^{-1},t)=1+d_{1k}(t)q^{-1}+d_{2k}(t)q^{-1}+\cdots+d_{n_dk}(t)q^{-n_d} \tag{4}$$

k是迭代次数，$u_k(t)$和$y_k(t)$是系统第k次运行时的输出，其中哈默斯坦静态非线性函数$f(u_k(t),t)$可以用不同的函数结构来建模，如神经网络结构、多项式函数、分段函数、饱和函数、死区函数等，本文把其看作已知基函数的多项式线性组合，$v_k(t)$是第k次运行时0均值，方差为σ^2的白噪声序列。根据输入信号$u_k(t)$映射到不可测的内部变量$\bar{u}_k(t)$[25]

$$\begin{aligned}\bar{u}_k(t)=&f(u_k(t),t)=h_{1k}(t)f_1(u_k(t))\\&+h_{2k}(t)f_2(u_k(t))+\cdots+h_{nk}(t)f_n(u_k(t))\\=&\sum_{l=1}^{n}h_{lk}(t)f_l(u_k(t))\end{aligned} \tag{5}$$

本文通过可测的输入输出数据$u_k(t)$, $y_k(t)$分析不同辨识算法对非线性模型参数$a_{ik}(t)$, $b_{ik}(t)$, $d_{ik}(t)$和$h_{ik}(t)$的估计效果，进一步比较系统的跟踪性能。下面将式(1)进一步展开，可以得到

$$\begin{aligned}y_k(t)=&[1-A(q^{-1},t)]y_k(t)+B(q^{-1},t)\bar{u}_k(t)\\&+D(q^{-1},t)v_k(t)\\=&-\sum_{i=1}^{n_a}a_{ik}(t)y_k(t-i)+\sum_{i=1}^{n_b}b_{ik}(t)\bar{u}_k(t-i)\\&+D(q^{-1},t)v_k(t)\end{aligned} \tag{6}$$

令

$$w_k(t)=D(q^{-1},t)v_k(t) \tag{7}$$

$$w_k(t)=\sum_{i=1}^{n_d}d_{ik}(t)v_k(t-i)+v_k(t) \tag{8}$$

可得

$$\begin{aligned}y_k(t)=&-\sum_{i=1}^{n_a}a_{ik}(t)y_k(t-i)+\sum_{i=1}^{n_b}b_{ik}(t)\bar{u}_k(t-i)\\&+\sum_{i=1}^{n_d}d_{ik}(t)v_k(t-i)+v_k(t)\end{aligned} \tag{9}$$

定义参数向量

$$\begin{aligned}\boldsymbol{\theta}_{ks}(t)=&[a_{1k}(t),\cdots,a_{n_ak}(t),h_{1k}(t)\boldsymbol{b}(t),\cdots,h_{nk}(t)\boldsymbol{b}(t)]^{\rm T}\\&\in\mathbb{R}^{n_1},n_1=n_a+n\end{aligned} \tag{10}$$

$$\begin{aligned}\boldsymbol{\theta}_k(t)=&[a_{1k}(t),\cdots,a_{n_ak}(t),h_{1k}(t)\boldsymbol{b}(t),\cdots,h_{nk}(t)\boldsymbol{b}(t),\\&d_{1k}(t),\cdots,d_{n_dk}(t)]^{\rm T}\in\mathbb{R}^{n_2},\\&n_2=n_a+n\times n_b+n_d\end{aligned}\tag{11}$$

同时，定义信息向量为

$$\begin{aligned}\boldsymbol{\phi}_{ks}(t)=&[-y_k(t-1),\cdots,-y_k(t-n_a),\varphi_1(t),\cdots,\varphi_n(t)]^{\rm T}\\&\in\mathbb{R}^{n_1}\end{aligned}\tag{12}$$

$$\begin{aligned}\boldsymbol{\phi}_k(t)=&[-y_k(t-1),\cdots,-y_k(t-n_a),\boldsymbol{\varphi}_1(t),\cdots,\boldsymbol{\varphi}_n(t),\\&v_k(t-1),\cdots,v_k(t-n_d)]^{\rm T}\in\mathbb{R}^{n_2}\end{aligned}\tag{13}$$

其中

$$\boldsymbol{b}(t)=[b_{1k}(t),\cdots,b_{n_bk}(t)]^{\rm T}\tag{14}$$

$$\boldsymbol{\varphi}_i(t)=[f_i(u_k(t-1)),\cdots,f_i(u_k(t-n_b))]^{\rm T}\tag{15}$$

从而获得哈默斯坦非线性时变系统的回归模型

$$y_k(t)=\boldsymbol{\phi}_k^{\rm T}(t)\boldsymbol{\theta}_k(t)+v_k(t)\tag{16}$$

由此可见，所构建的模型形成了可沿时间轴递推和重复轴学习的回归形式，可采用基于重复轴的迭代学习算法获得参数的一致估计。由于式(2)中系统参数$h_{ik}(t)\boldsymbol{b}(t)$存在耦合现象，需对其系数进行归一化处理。令$\boldsymbol{\theta}_{lk}^i(t)=h_{lk}(t)b_{ik}(t)$，通过辨识算法得到$\boldsymbol{\theta}_{lk}^i(t)$，假设$h_{lk}(t)$的第1项$h_{1k}(t)=1$，可以得到$b_{ik}(t)$，再对其做除法和算术平均处理，从而获得$h_{lk}(t)=\sum_{i=1}^{n_b}\dfrac{\boldsymbol{\theta}_{lk}^i(t)}{b_{ik}(t)}/n_b,\ l=2,3,\cdots,n$。

3 哈默斯坦非线性时变系统的加权学习辨识

当系统在给定区间上重复运行时，记录或量测从第1次到第k次操作的输入和输出数据，对应固定时刻t

$$\left.\begin{aligned}\boldsymbol{Y}_k(t)&=[y_1(t),y_2(t),\cdots,y_k(t)]^{\rm T}\\\boldsymbol{\Phi}_k(t)&=[\phi_1(t),\phi_2(t),\cdots,\phi_k(t)]^{\rm T}\\\boldsymbol{V}_k(t)&=[v_1(t),v_2(t),\cdots,v_k(t)]^{\rm T}\end{aligned}\right\}\tag{17}$$

可将表示k次重复运行的系统特性表达为

$$\boldsymbol{Y}_k(t)=\boldsymbol{\Phi}_k^{\rm T}(t)\boldsymbol{\theta}_k(t)+\boldsymbol{V}_k(t)\tag{18}$$

3.1 迭代学习最小二乘算法

针对重复运行的哈默斯坦时变系统，考虑如式(19)的极小化准则函数

$$\begin{aligned}J_k(\hat{\boldsymbol{\theta}}_k(t),t)=&\frac{1}{2}[\boldsymbol{Y}_k(t)-\boldsymbol{\Phi}_k(t)\hat{\boldsymbol{\theta}}_k(t)]^{\rm T}\\&\cdot[\boldsymbol{Y}_k(t)-\boldsymbol{\Phi}_k(t)\hat{\boldsymbol{\theta}}_k(t)]\end{aligned}\tag{19}$$

假设$\boldsymbol{\Phi}_k^{\rm T}(t)\boldsymbol{\Phi}_k(t)$是可逆的，欲使式(19)最小，可得

$$\hat{\boldsymbol{\theta}}_k(t)=(\boldsymbol{\Phi}_k^{\rm T}(t)\boldsymbol{\Phi}_k(t))^{-1}\boldsymbol{\Phi}_k^{\rm T}(t)\boldsymbol{Y}_k(t)\tag{20}$$

由于式(11)中$\boldsymbol{\Phi}_k(t)[\boldsymbol{\phi}_k(t)]$包含了未知项$v_k(t-i)$，考虑辅助模型的方法，借助交互估计理论的思想，对于信息矩阵中的未知项$v_k(t-i)$用其估计$\hat{v}_k(t-i)$代替，则$\boldsymbol{\phi}_k(t)$可用$\hat{\boldsymbol{\phi}}_{\boldsymbol{k}}(t)$代替，可得

$$\begin{aligned}\hat{\boldsymbol{\phi}}_k(t)=&[-y_k(t-1),\cdots,-y_k(t-n_a),\boldsymbol{\varphi}_1(t),\cdots,\\&\boldsymbol{\varphi}_n(t),\hat{v}_k(t-1),\cdots,\hat{v}_k(t-n_d)]^{\rm T}\end{aligned}\tag{21}$$

由式(16)可知

$$v_k(t)=y_k(t)-\boldsymbol{\phi}_k^{\rm T}(t)\boldsymbol{\theta}_k(t)\tag{22}$$

$$\hat{v}_k(t)=y_k(t)-\hat{\boldsymbol{\phi}}_k^{\rm T}(t)\hat{\boldsymbol{\theta}}_k(t)\tag{23}$$

当采集到系统重复运行的一批数据后，通过上式可以获得时变参数的估计，但当$\hat{\boldsymbol{\theta}}_k(t)$维数较大时，求逆运算将使计算量增大，为了回避求逆运算，定义

$$\boldsymbol{P}_k{}^{-1}(t)=\boldsymbol{\Phi}_k^{\rm T}(t)\boldsymbol{\Phi}_k(t)\tag{24}$$

可得

$$\boldsymbol{P}_k^{-1}(t)=\boldsymbol{P}_{k-1}^{-1}(t)+\hat{\boldsymbol{\phi}}_k(t)\hat{\boldsymbol{\phi}}_k^{\rm T}(t)\tag{25}$$

利用矩阵求逆公式，$(\boldsymbol{A}+\boldsymbol{BCD})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1}+\boldsymbol{DA}^{-1}\boldsymbol{B}^{-1})^{-1}\boldsymbol{DA}^{-1}$，可得

$$\boldsymbol{P}_k(t)=\boldsymbol{P}_{k-1}(t)-\frac{\boldsymbol{P}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)\hat{\boldsymbol{\phi}}_k^{\rm T}(t)\boldsymbol{P}_{k-1}(t)}{1+\hat{\boldsymbol{\phi}}_k^{\rm T}(t)\boldsymbol{P}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)}\tag{26}$$

式(26)两端同乘$\hat{\boldsymbol{\phi}}_k(t)$

$$\boldsymbol{P}_k(t)\hat{\boldsymbol{\phi}}_k(t)=\frac{\boldsymbol{P}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)}{1+\hat{\boldsymbol{\phi}}_k^{\rm T}(t)P_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)}\tag{27}$$

将式(25)代入式(20)，可得

$$\begin{aligned}\hat{\boldsymbol{\theta}}_k(t)=&\boldsymbol{P}_k(t)\boldsymbol{\Phi}_k^{\rm T}(t)Y_k(t)\\=&\boldsymbol{P}_k(t)[\boldsymbol{\Phi}_{k-1}^{\rm T}(t)\boldsymbol{Y}_{k-1}(t)+\hat{\boldsymbol{\phi}}_k(t)y_k(t)]\\=&\boldsymbol{P}_k(t)[\boldsymbol{P}_{k-1}^{-1}(t)\hat{\boldsymbol{\theta}}_{k-1}(t)+\hat{\boldsymbol{\phi}}_k(t)y_k(t)]\end{aligned}\tag{28}$$

根据式(25)、式(26)和式(28)，可得

$$\begin{aligned}\hat{\boldsymbol{\theta}}_k(t)=&\hat{\boldsymbol{\theta}}_{k-1}(t)+\frac{\boldsymbol{P}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)}{1+\hat{\boldsymbol{\phi}}_k^{\rm T}(t)\boldsymbol{P}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)}\\&\cdot[y_k(t)-\hat{\boldsymbol{\phi}}_k^{\rm T}(t)\hat{\boldsymbol{\theta}}_{k-1}(t)]\end{aligned}\tag{29}$$

至此，式(26)和式(29)构成了该模型的迭代学习最小二乘算法。

3.2 带遗忘因子的迭代学习最小二乘算法

为避免数据饱和现象，增加新数据的权重，加快跟踪误差收敛速度及精度，在上述算法的基础上引入遗忘因子λ，考虑如式(30)的准则函数

$$\begin{aligned}J_k(\hat{\boldsymbol{\theta}}_k(t),t)=&\frac{1}{2}[\boldsymbol{Y}_k(t)-\boldsymbol{\Phi}_k(t)\hat{\boldsymbol{\theta}}_k(t)]^{\rm T}\\&\cdot\boldsymbol{L}_k[\boldsymbol{Y}_k(t)-\boldsymbol{\Phi}_k(t)\hat{\boldsymbol{\theta}}_k(t)]\end{aligned}\tag{30}$$

其中

$$\boldsymbol{L}_k=\begin{bmatrix}\lambda^{k-1} & & & \\ & \lambda^{k-2} & & \\ & & \ddots & \\ & & & 1\end{bmatrix}_{k\times k} \tag{31}$$

欲使式(30)最小，令$\dfrac{\partial J_k(\hat{\boldsymbol{\theta}}_k(t),t)}{\hat{\boldsymbol{\theta}}_k(t)}=0$

$$\frac{\partial J_k(\hat{\boldsymbol{\theta}}_k(t),t)}{\hat{\boldsymbol{\theta}}_k(t)}=\boldsymbol{\Phi}_k^{\mathrm{T}}(t)\boldsymbol{L}_k\boldsymbol{\Phi}_k(t)\hat{\boldsymbol{\theta}}_k(t)-\boldsymbol{\Phi}_k^{\mathrm{T}}(t)\boldsymbol{L}_k\boldsymbol{Y}_k(t)=0 \tag{32}$$

可得

$$\hat{\boldsymbol{\theta}}_k(t)=(\boldsymbol{\Phi}_k^{\mathrm{T}}(t)\boldsymbol{L}_k\boldsymbol{\Phi}_k(t))^{-1}\boldsymbol{\Phi}_k^{\mathrm{T}}(t)\boldsymbol{L}_k\boldsymbol{Y}_k(t) \tag{33}$$

假设$\boldsymbol{\Phi}_k^{\mathrm{T}}(t)\boldsymbol{L}_k\boldsymbol{\Phi}_k(t)$是可逆的，参照式(24)—式(29)，定义

$$\boldsymbol{H}_k^{-1}(t)=\boldsymbol{\Phi}_k^{\mathrm{T}}(t)\boldsymbol{L}_k\boldsymbol{\Phi}_k(t) \tag{34}$$

$$\boldsymbol{H}_k^{-1}(t)=\lambda\boldsymbol{H}_{k-1}^{-1}(t)+\hat{\boldsymbol{\phi}}_k(t)\hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(t) \tag{35}$$

$$\lambda\boldsymbol{H}_k(t)=\boldsymbol{H}_{k-1}(t)-\frac{\boldsymbol{H}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)\hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(t)\boldsymbol{H}_{k-1}(t)}{\lambda+\hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(t)\boldsymbol{H}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)} \tag{36}$$

两端同乘$\hat{\boldsymbol{\phi}}_k(t)$，除以$\lambda$可得

$$\boldsymbol{H}_k(t)\hat{\boldsymbol{\phi}}_k(t)=\frac{\boldsymbol{H}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)}{\lambda+\hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(t)\boldsymbol{H}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)} \tag{37}$$

$$\begin{aligned}\hat{\boldsymbol{\theta}}_k(t)=&\boldsymbol{H}_k(t)\boldsymbol{\Phi}_k^{\mathrm{T}}(t)\boldsymbol{L}_k(t)\boldsymbol{Y}_k(t)\\=&\boldsymbol{H}_k(t)[\lambda\boldsymbol{\Phi}_{k-1}^{\mathrm{T}}(t)\boldsymbol{L}_{k-1}\boldsymbol{Y}_{k-1}(t)+\hat{\boldsymbol{\phi}}_k(t)y_k(t)]\\=&\boldsymbol{H}_k(t)[\lambda\boldsymbol{H}_{k-1}^{-1}(t)\hat{\boldsymbol{\theta}}_{k-1}(t)+\hat{\boldsymbol{\phi}}_k(t)y_k(t)]\\=&\boldsymbol{H}_k(t)[(\boldsymbol{H}_k^{-1}(t)-\hat{\boldsymbol{\phi}}_k(t)\hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(t))\hat{\boldsymbol{\theta}}_{k-1}(t)\\&+\hat{\boldsymbol{\phi}}_k(t)y_k(t)]\\=&\hat{\boldsymbol{\theta}}_{k-1}(t)+\boldsymbol{H}_k(t)\hat{\boldsymbol{\phi}}_k(t)(y_k(t)-\hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{k-1}(t))\end{aligned} \tag{38}$$

把式(37)代入

$$\begin{aligned}\hat{\boldsymbol{\theta}}_k(t)=&\hat{\boldsymbol{\theta}}_{k-1}(t)+\frac{\boldsymbol{H}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)}{\lambda+\hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(t)\boldsymbol{H}_{k-1}(t)\hat{\boldsymbol{\phi}}_k(t)}\\&\cdot(y_k(t)-\hat{\boldsymbol{\phi}}_k^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{k-1}(t))\end{aligned} \tag{39}$$

式(37)和式(39)构成带遗忘因子迭代学习最小二乘算法。

3.3 加权迭代学习最小二乘算法

上述两类算法能获得参数的一致估计，但对初始条件的一致性要求苛刻，且跟踪精度有进一步提升的空间，下面提出一种加权的迭代学习算法以获得更优跟踪效果和辨识精度。为了得到参数$\boldsymbol{\theta}_k(t)$的估计，考虑如式(40)的准则函数[21]

$$\begin{aligned}\boldsymbol{J}_k(\hat{\boldsymbol{\theta}}_k(t),t)=&\boldsymbol{E}_k{}^{\mathrm{T}}(t)\boldsymbol{Q}_k\boldsymbol{E}_k(t)+(\hat{\boldsymbol{\theta}}_k(t)-\hat{\boldsymbol{\theta}}_{k-1}(t))^{\mathrm{T}}\\&\cdot\boldsymbol{R}_k(\hat{\boldsymbol{\theta}}_k(t)-\hat{\boldsymbol{\theta}}_{k-1}(t))\end{aligned} \tag{40}$$

其中

$$\boldsymbol{E}_k(t)=\boldsymbol{Y}_k(t)-\boldsymbol{\Phi}_k^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_k(t) \tag{41}$$

$\boldsymbol{E}_k(t)$为系统模型输出误差构造成的矩阵，$\boldsymbol{Q}_k$和$\boldsymbol{R}_k$为正定的权重矩阵，这里，$q_{k_n}(t)=q_1^{k_n}$，$k_n\in[1,k]$,$q_1\in[0,1]$,$r_1\in[0,1]$ 。

$$\boldsymbol{Q}_k=\begin{bmatrix}q_1 & \cdots & 0 & \cdots & 0\\ \vdots & \ddots & \vdots & & \vdots\\ 0 & \cdots & q_{k_n} & \cdots & 0\\ \vdots & & \vdots & \ddots & \vdots\\ 0 & \cdots & 0 & \cdots & q_k\end{bmatrix}_{k\times k} \tag{42}$$

$$\boldsymbol{R}_k=\begin{bmatrix}r_1^{n_2-1} & & & \\ & r_1^{n_2-2} & & \\ & & \ddots & \\ & & & 1\end{bmatrix}_{n_2\times n_2} \tag{43}$$

根据矩阵求导公式

$$\begin{aligned}\frac{\partial J_k(\hat{\boldsymbol{\theta}}_k(t),t)}{\boldsymbol{\theta}_k(t)}=&-2\boldsymbol{\Phi}_k^{\mathrm{T}}(t)Q_k(Y_k(t)-\boldsymbol{\Phi}_k^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_k(t))\\&+2R_k(\hat{\boldsymbol{\theta}}_k(t)-\hat{\boldsymbol{\theta}}_{k-1}(t))\end{aligned} \tag{44}$$

欲使式(44)最小，可令

$$\frac{\partial J_k(\hat{\boldsymbol{\theta}}_k(t),t)}{\hat{\boldsymbol{\theta}}_k(t)}=0 \tag{45}$$

此时可获得系统模型输出误差与参数估计误差的最优化，求得

$$\begin{aligned}\hat{\boldsymbol{\theta}}_k(t)=&(\boldsymbol{\Phi}_k^{\mathrm{T}}(t)Q_k\boldsymbol{\Phi}_k(t)+R_k)^{-1}R_k\hat{\boldsymbol{\theta}}_{k-1}(t)\\&+(\boldsymbol{\Phi}_k^{\mathrm{T}}(t)Q_k\boldsymbol{\Phi}_k(t)+R_k)^{-1}\boldsymbol{\Phi}_k^{\mathrm{T}}(t)Q_kY_k(t)\end{aligned} \tag{46}$$

从而，式(46)构成该模型加权最小二乘算法，通过加权最小二乘算法进行哈默斯坦模型时变参数估计的步骤如表1所示。

4　数值算例

本节将完成具体的算例仿真，考虑如下重复运行的时变哈默斯坦模型

$$A(q^{-1},t)y_k(t)=B(q^{-1},t)f(u_k(t),t)+D(q^{-1},t)v_k(t) \tag{47}$$

k是迭代次数，$u_k(t)$和$y_k(t)$是系统第k次运行时的输出，其中哈默斯坦静态非线性函数$f(u_k(t),t)$可以用不同的函数结构来建模，$v_k(t)$是第k次运行时零均值、方差为σ^2的白噪声序列，输入信号$u_k(t)$映射到不可测的非线性内部变量$f(u_k(t),t)$。线性部分参数$A(q^{-1},t)=1+a_{1k}(t)q^{-1}+a_{2k}(t)q^{-1}$，$B(q^{-1},t)=b_{1k}(t)q^{-1}+b_{2k}(t)q^{-1}$，$D(q^{-1},t)=1+d_{1k}(t)q^{-1}$。其中，$a_{1k}(t)=1.5^{\sin\frac{1000}{t}}-1$，$a_{2k}(t)=0.01\times t\times\lg\dfrac{50}{t}$，$b_{1k}(t)=0.16\times\sin\dfrac{3\pi t}{20}-1.2$，$b_{2k}(t)=\lg\Big(0.1+1.4\times$

表 1 采用加权迭代学习最小二乘算法进行参数估计流程图

输入： 重复激励的一组数列

输出： 堆积的输出向量$\boldsymbol{Y}_k(t)$

(1) 对于所有的$t=0,1,\cdots,N$，给定参数估计的初始值$\hat{\theta}_{-1}(t)$=0，迭代所需的$\hat{v}_k(t)$,q_1及r_1，并置$k=0$;

(2) While $k \le K_{\max}$($K_{\max}$为最大迭代次数)

(3) for each $t \in [0,N]$

(4) 在第k次重复运行时，采集输入数据$u_k(t)$，计算输出数据$y_k(t)$;

(5) 计算$\boldsymbol{Q}_k$和$\boldsymbol{\Phi}_k(t)[\hat{\boldsymbol{\phi}}_k(t)]$;

(6) 通过式(46)计算得出$\hat{\boldsymbol{\theta}}_k(t)$;

(7) 利用式(23)更新$\hat{v}_k(t)$;

(8) End

(9) 检验迭代停止条件，满足则停止；否则置$k=k+1$，并回到第3步；

(10) End

$\left|\cos\dfrac{30\pi}{t}\right|\Big)-1$, $d_{1k}(t)=0.17\times\sin\dfrac{2\pi t}{15}+0.3$。非线性部分，$f(u_k(t),t)=h_{1k}(t)u_k(t)+h_{2k}(t)u_k^2(t)+h_{3k}(t)u_k^3(t)$，为了避免参数耦合问题，假设$h_{1k}(t)=1$，其余参数为$h_{2k}(t)=0.7+1.3\times\left|\sin\dfrac{36\pi}{t}\right|^{\frac{1}{8}}$，$h_{3k}(t)=0.05\times\cos\dfrac{3\pi t}{10}-0.9$。

在仿真过程中，取输入$u_k(t)$为零均值、方差为1的不相关持续激励信号，噪声项为零均值、方差为0.01的白噪声序列。选取初值$P_{-1}(t)=p_{-1}(t)I_{9\times 9}$，$p_{-1}(t)=10^6$，$\boldsymbol{\theta}_{-1}=[0,0,0,0,0,0,0,0,0]^{\mathrm{T}}$，$r_1=0.98$, $q_1=0.98$，分别采用迭代学习最小二乘算法、带遗忘因子迭代学习最小二乘算法，以及加权迭代学习最小二乘算法估计该时变系统的参数。采用带遗忘因子迭代学习最小二乘算法的参数估计中，遗忘因子取$\lambda=0.98$，为检验其收敛效果，定

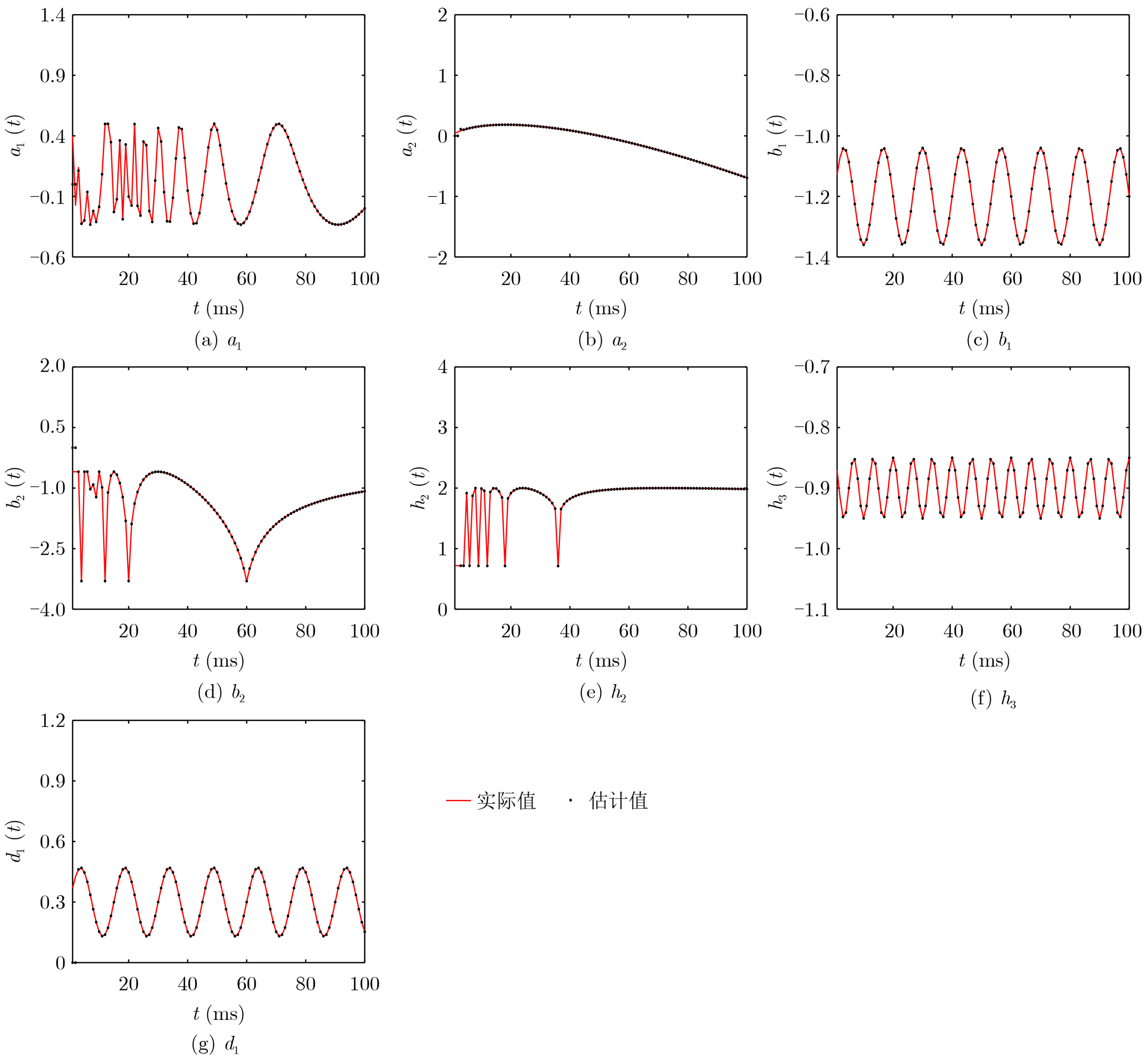
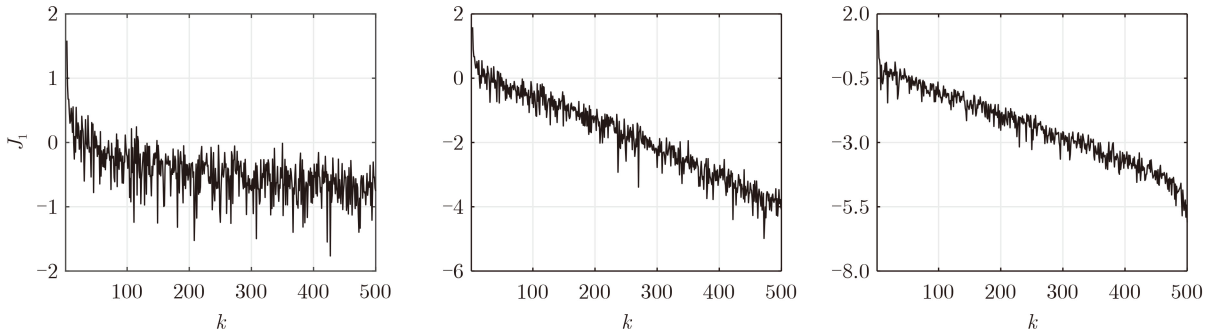


图 2 采用加权迭代学习最小二乘算法的参数估计结果

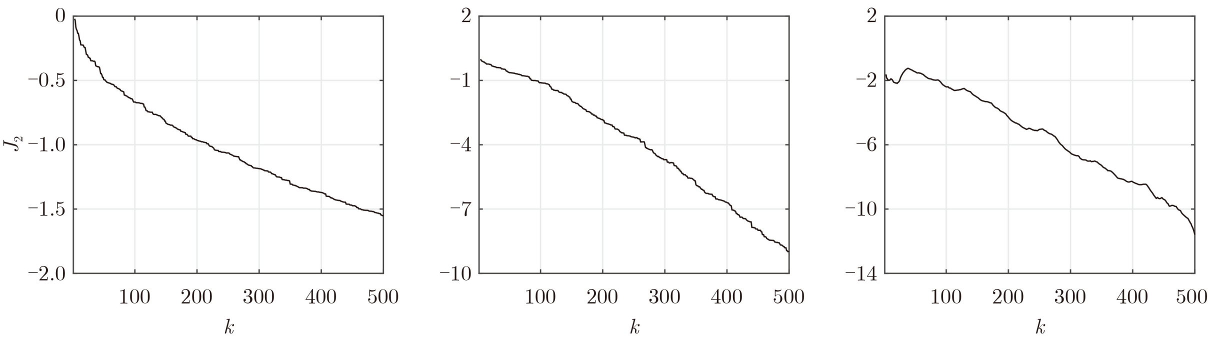
义系统每次迭代的模型输出误差$J_1 = \ln \max\limits_{0\le t\le N} |e_k(t)|$，这里$e_k(t) = y_k(t) - \hat{\varphi}_k^{\mathrm{T}}(t)\hat{\theta}_{k-1}(t)$，定义系统每次迭代的参数估计误差$J_2 = \ln \max\limits_{0\le t\le N} |\delta_k(t)|$，这里$\delta_k(t) = \dfrac{||\hat{\theta}_k(t) - \theta_{-1}(t)||}{||\theta_{-1}(t)||}$。

加权迭代学习算法的参数估计结果如图2所示，3种算法的模型输出误差和参数估计误差如图3和图4所示。仿真结果表明，加权迭代学习算法可以有效估计模型的时变参数。在迭代学习最小二乘算法的基础上引入遗忘因子可以改善模型输出误差和参数估计误差。考虑如式(40)的准则函数，引入加权矩阵，可以进一步降低模型输出误差，提高辨识精度，尤其在达到某一确定的参数估计误差或模型输出误差时，迭代次数明显减少，同时该算法可获得更好的辨识效果。



(a) 迭代学习最小二乘算法模型输出误差　(b) 带遗忘因子迭代学习最小二乘算法模型输出误差　(c) 加权迭代学习最小二乘算法模型输出误差

图 3 采用3种不同算法的模型输出误差比较



(a) 迭代学习最小二乘算法参数估计误差　(b) 带遗忘因子迭代学习最小二乘算法参数估计误差　(c) 加权迭代学习最小二乘算法参数估计误差

图 4 采用3种算法的参数估计误差比较

5　结论

对于Harmmerstein非线性时变系统的参数估计问题，本文提出加权学习辨识算法，推导了重复运行条件下的迭代学习最小二乘法、带遗忘因子迭代学习最小二乘法和加权迭代学习最小二乘法。在重复持续激励条件下，仿真结果显示加权学习辨识算法可以实现时变参数的完全估计。与现有辨识算法相比，本文算法进一步提高了时变参数的估计精度，提高了算法的跟踪性能，在达到一定辨识精度下，加权算法可以减少重复运行的次数。

参 考 文 献

[1] WANG Dongqing, ZHANG Shuo, GAN Min, *et al*. A novel EM identification method for hammerstein systems with missing output data[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(4): 2500–2508. doi: 10.1109/TII.2019.2931792.

[2] CERONE V, RAZZA V, and REGRUTO D. One-shot set-membership identification of Generalized Hammerstein–Wiener systems[J]. *Automatica*, 2020, 118: 109028. doi: 10.1016/j.automatica.2020.109028.

[3] LYU Bensheng, LI Jia, and LI Feng. Neuro-fuzzy based identification of Hammerstein OEAR systems[J]. *Computers & Chemical Engineering*, 2020, 141: 106984. doi: 10.1016/j.compchemeng.2020.106984.

[4] 贾立, 李训龙. Hammerstein模型辨识的回顾及展望[J]. 控制理论与应用, 2014, 31(1): 1–10. doi: 10.7641/CTA.2014.30478.

JIA Li and LI Xunlong. Identification of Hammerstein model: Review and prospect[J]. *Control Theory & Applications*, 2014, 31(1): 1–10. doi: 10.7641/CTA.2014.30478.

[5] WESTWICK D and VERHAEGEN M. Identifying MIMO Wiener systems using subspace model identification methods[J]. *Signal Processing*, 1996, 52(2): 235–258. doi: 10.1016/0165-1684(96)00056-4.

[6] LOVERA M, GUSTAFSSON T, and VERHAEGEN M. Recursive subspace identification of linear and non-linear Wiener state-space models[J]. *Automatica*, 2000, 36(11): 1639–1650. doi: 10.1016/S0005-1098(00)00103-5.

[7] JALALEDDINI K and KEARNEY R E. Subspace identification of SISO hammerstein systems: Application to stretch reflex identification[J]. *IEEE Transactions on Biomedical Engineering*, 2013, 60(10): 2725–2734. doi: 10.1109/TBME.2013.2264216.

[8] GOETHALS I, PELCKMANS K, SUYKENS J A K, *et al*. Subspace identification of hammerstein systems using least squares support vector machines[J]. *IEEE Transactions on Automatic Control*, 2005, 50(10): 1509–1519. doi: 10.1109/TAC.2005.856647.

[9] BAI Erwei. An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems[J]. *Automatica*, 1998, 34(3): 333–338. doi: 10.1016/S0005-1098(97)00198-2.

[10] CHANG F H I and LUUS R. A noniterative method for identification using Hammerstein model[J]. *IEEE Transactions on Automatic Control*, 1971, 16(5): 464–468. doi: 10.1109/TAC.1971.1099787.

[11] WANG Dongqing and DING Feng. Least squares based and gradient based iterative identification for Wiener nonlinear systems[J]. *Signal Processing*, 2011, 91(5): 1182–1189. doi: 10.1016/j.sigpro.2010.11.004.

[12] WANG Dongqing, FAN Qiuhua, and MA Yan. An interactive maximum likelihood estimation method for multivariable Hammerstein systems[J]. *Journal of the Franklin Institute*, 2020, 357(17): 12986–13005. doi: 10.1016/j.jfranklin.2020.09.005.

[13] GREBLICKI W and PAWLAK M. The weighted nearest neighbor estimate for Hammerstein system identification[J]. *IEEE Transactions on Automatic Control*, 2019, 64(4): 1550–1565. doi: 10.1109/TAC.2018.2866463.

[14] MZYK G and WACHEL P. Kernel-based identification of Wiener–Hammerstein system[J]. *Automatica*, 2017, 83: 275–281. doi: 10.1016/j.automatica.2017.06.038.

[15] GIORDANO G, GROS S, and SJÖBERG J. An improved method for Wiener-Hammerstein system identification based on the Fractional Approach[J]. *Automatica*, 2018, 94: 349–360. doi: 10.1016/j.automatica.2018.04.046.

[16] SUNG S W. System identification method for Hammerstein processes[J]. *Industrial & Engineering Chemistry Research*, 2002, 41(17): 4295–4302. doi: 10.1021/ie0109206.

[17] JIA Li, LI Xunlong, and CHIU M S. Correlation analysis based MIMO neuro-fuzzy Hammerstein model with noises[J]. *Journal of Process Control*, 2016, 41: 76–91. doi: 10.1016/j.jprocont.2015.11.006.

[18] DING Feng, SHI Yang, and CHEN Tongwen. Gradient-based identification methods for hammerstein nonlinear ARMAX models[J]. *Nonlinear Dynamics*, 2006, 45(1/2): 31–43. doi: 10.1007/s11071-005-1850-z.

[19] REN Biying, XIE Chenxue, SUN Xiangdong, *et al*. Parameter identification of a lithium-ion battery based on the improved recursive least square algorithm[J]. *IET Power Electronics*, 2020, 13(12): 2531–2537. doi: 10.1049/iet-pel.2019.1589.

[20] ZHANG Bo, TANG Yinggan, and LU Yao. Identification of linear time-varying fractional order systems using block pulse functions based on repetitive principle[J]. *ISA Transactions*, 2022, 123: 218–229. doi: 10.1016/j.isatra.2021.05.024..

[21] 孙明轩, 毕宏博. 学习辨识: 最小二乘算法及其重复一致性[J]. 自动化学报, 2012, 38(5): 698–706. doi: 10.3724/SP.J.1004.2012.00698.

SUN Mingxuan and BI Hongbo. Learning identification: Least squares algorithms and their repetitive consistency[J]. *Acta Automatica Sinica*, 2012, 38(5): 698–706. doi: 10.3724/SP.J.1004.2012.00698.

[22] SONG Fazhi, LIU Yang, WANG Xianli, *et al*. Enhancing accuracy and numerical stability for repetitive time-varying system identification: An iterative learning approach[J]. *IEEE Access*, 2020, 8: 25679–25690. doi: 10.1109/ACCESS.2020.2966300.

[23] LIU Nanjun and ALLEYNE A. Iterative learning identification for linear time-varying systems[J]. *IEEE Transactions on Control Systems Technology*, 2016, 24(1): 310–317. doi: 10.1109/TCST.2015.2424374.

[24] DING Feng, XU Ling, MENG Dandan, *et al*. Gradient estimation algorithms for the parameter identification of bilinear systems using the auxiliary model[J]. *Journal of Computational and Applied Mathematics*, 2020, 369: 112575. doi: 10.1016/j.cam.2019.112575.

[25] DING Jie, CAO Zhengxin, CHEN Jiazhong, *et al*. Weighted parameter estimation for hammerstein nonlinear ARX systems[J]. *Circuits, Systems, and Signal Processing*, 2020, 39(4): 2178–2192. doi: 10.1007/s00034-019-01261-4.

仲国民：男，1983年生，博士生，研究方向为系统辨识与学习控制.

俞其乐：男，1997年生，硕士生，研究方向为学习辨识.

陈　强：男，1984年生，副教授，硕士生导师，主要研究方向为自适应与学习控制.

责任编辑：余　蓉