

# 控制与决策

Control and Decision

有理模型辨识的两类新方法——混合迭代与柔性最小二乘法

陈晶, 朱全民

引用本文:

陈晶, 朱全民. 有理模型辨识的两类新方法——混合迭代与柔性最小二乘法[J]. 控制与决策, 2022, 37(1): 58–66.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2020.0935>

---

## 您可能感兴趣的其他文章

Articles you may be interested in

### 多策略融合的改进麻雀搜索算法及其应用

Improved sparrow search algorithm with multi-strategy integration and its application

控制与决策. 2022, 37(1): 87–96 <https://doi.org/10.13195/j.kzyjc.2021.0582>

### 分布式最小二乘估计中隐匿FDI攻击策略的设计

Hidden FDI attack strategy for distributed least square estimation

控制与决策. 2021, 36(8): 1963–1969 <https://doi.org/10.13195/j.kzyjc.2019.1688>

### 基于多维泰勒网的超前d步预测模型

d-step-ahead predictive model based on multi-dimensional Taylor network

控制与决策. 2021, 36(2): 345–354 <https://doi.org/10.13195/j.kzyjc.2019.0722>

### 基于KRLS的非均匀采样非线性系统辨识

Identification of non-uniformly sampled nonlinear systems based on KRLS

控制与决策. 2021, 36(12): 3049–3055 <https://doi.org/10.13195/j.kzyjc.2020.0962>

### 基于高斯混合模型聚类的非均匀采样系统的多模型切换辨识

Multi-model switching identification for non-uniformly sampled systems based on Gaussian mixture model clustering

控制与决策. 2021, 36(12): 2946–2954 <https://doi.org/10.13195/j.kzyjc.2020.0671>

# 有理模型辨识的两类新方法——混合迭代与柔性最小二乘法

陈晶<sup>1†</sup>, 朱全民<sup>2</sup>

(1. 江南大学理学院, 江苏无锡 214122; 2. 西英格兰大学工业设计和数学系, 英国布里斯托 BS16 1QY)

**摘要:** 针对有理模型提出两类辨识方法. 首先提出基于递阶辨识思想的混合辨识方法, 将模型分解为分子和分母两个子模型, 分别用最小二乘法辨识分子参数, 用粒子群算法和智能多步长梯度迭代算法辨识分母参数. 由于降低了模型维数, 且信息向量与噪声不相关, 相对于传统的偏差补偿最小二乘法, 混合迭代法可以提高辨识精度并降低计算量. 然后, 为消除模型结构已知的假设, 且充分利用最新数据更新系统参数, 提出柔性递推最小二乘法, 将有理模型转化为时变参数系统, 进而辨识出时变系统的参数. 仿真例子验证了所提出方法的有效性.

**关键词:** 有理模型; 参数估计; 梯度迭代; 粒子群算法; 柔性递推最小二乘法; 混合辨识算法

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2020.0935

开放科学(资源服务)标识码(OSID):



**引用格式:** 陈晶, 朱全民. 有理模型辨识的两类新方法——混合迭代与柔性最小二乘法[J]. 控制与决策, 2022, 37(1): 58-66.

## Two novel identification methods for rational models—Compound iterative algorithm and flexible least squares algorithm

CHEN Jing<sup>1†</sup>, ZHU Quan-min<sup>2</sup>

(1. School of Science, Jiangnan University, Wuxi 214122, China; 2. Department of Engineering Design and Mathematics, University of the West of England, Bristol BS16 1QY, U.K.)

**Abstract:** This study proposes two identification methods for nonlinear rational models. The first is the compound iterative algorithm which is based on the hierarchical technique, which transforms the rational model into two sub-models whose parameters are estimated iteratively using the particle swarm optimization algorithm and intelligent multi-step-length algorithm, respectively. Thus, it has less computational efforts and higher estimation accuracy when compared with the traditional identification methods. Then, a flexible recursive least squares algorithm is proposed which turns the rational model into a time-varying model, thus it does not require the knowledge of the structure of the denominator model, and can update the parameters with new collected data. Simulation results show the effectiveness of the proposed algorithms.

**Keywords:** rational model; parameter estimation; gradient iterative; particle swarm optimization algorithm; flexible recursive least squares algorithm; compound iterative algorithm

## 0 引言

有理模型 (rational model) 是由英国科学家 Billings 教授<sup>[1]</sup>提出的一种特殊的非线性模型, 其模型结构是用一个分式来描述. 它具有一般非线性模型所不具有的独特的特点: 1) 所有的多项式非线性模型都可以用有理模型描述; 2) 可以用一个简单的结构逼近一个复杂的多项式模型; 3) 分母较小的变化可能会引起输出的巨大跳变; 4) 工业过程中很多复杂的动态特性可由其进行建模, 如化工领域<sup>[2]</sup>、生

命科学以及经济领域<sup>[3-5]</sup>. 因此, 对有理模型的研究具有广泛的理论和应用意义.

要对被控对象实现精准控制或者对控制系统的故障进行检测和诊断, 被控对象/控制系统的模型需要事先已知<sup>[6-8]</sup>. 因此, 对有理模型进行辨识研究势在必行. 目前, 针对线性系统或者多项式非线性系统存在多种辨识方法, 如: 最小二乘辨识算法<sup>[9-10]</sup>、梯度类算法<sup>[11-12]</sup>、极大似然辨识算法<sup>[13]</sup>、最大期望辨识算法等<sup>[14-15]</sup>. 然而, 由于有理模型分母存在未知参数, 以上

收稿日期: 2020-07-10; 录用日期: 2020-11-05.

基金项目: 国家自然科学基金项目(61973137); 近地面探测技术重点实验室基金项目(TCGZ2019A001); 中央高校基本科研业务费专项资金项目(JUSRP22016).

责任编辑: 柴利.

<sup>†</sup>通讯作者. E-mail: chenjing1981929@126.com.

经典的辨识算法不能直接应用于有理模型辨识中。

针对有理模型的辨识,最常见的方法是通过等式转换将分式模型转化为线性自回归模型,进而对自回归模型进行辨识.文献[16]针对有理模型,先将有理模型转换为线性模型,因转化为线性模型后,信息向量中存在与噪声相关的项,故提出了基于偏差补偿的最小二乘算法.文献[17]利用神经网络结构来描述有理模型,进而提出了基于神经网络的BP算法.上述文献提出的都是离线辨识算法,由于离线辨识算法不能利用新收集数据来更新参数,且计算量大,文献[18]针对有理模型提出了递推最小二乘算法,每一次都只使用一个新的数据来进行参数更新,计算量大大减小.文献[19]借助递阶辨识思想,将有理模型分解为分子和分母两个子模型,对两个耦合的子模型分别使用递推最小二乘和反向递推最小二乘算法进行辨识.该方法将复杂的有理模型辨识转化为两个简单的子模型辨识,计算量大大降低,但是,辨识分母时要求系统的噪信比尽可能小,因而限制了其使用范围.

20世纪80年代末,美国统计学家Kalaba和Teshfatsion<sup>[20]</sup>提出利用柔性最小二乘算法(flexible least squares algorithm)辨识时变参数系统,并用来预测美国货币的需求情况.该算法的主要思想是设计一个关于测量误差和相连参数误差的损失函数,并为每一类误差设置不同的权值:当参数变化较快时,为参数误差设置较小的权值;相应地,当参数缓慢变化时,为其设置较大的权值<sup>[21-22]</sup>.利用损失函数分别对每一个时刻的参数求偏导并设置扩展的矩阵进而辨识出时变系统的参数.由于有理模型的特点,可以将包含输入和输出项的分母看成是一个时变参数,进而将有理模型转化为时变参数系统.

本文借助递阶辨识思想和柔性最小二乘思想,将有理模型转化为分子模型和分母模型,分别对分子模型和分母模型进行辨识.针对分母模型难以辨识的问题,提出粒子群算法;但由于粒子群算法计算量大且算法的稳定性密切依赖于粒子的数目,进一步提出智能多步长梯度迭代算法.该算法为每个负梯度方向设置多个步长,通过损失函数找出最优步长以及其对应的参数.上述基于递阶辨识思想的方法需要知道分母的结构,因而限制了其使用范围.为克服这个缺点,本文根据有理模型的特点,将有理模型的分母项和分子参数组成一个新的时变参数向量,并通过柔性递推最小二乘算法辨识出该时变参数向量.仿真例子验证了本文方法的有效性.

本文的主要贡献如下:

1)有理模型因其特殊结构,其辨识难度较大,借助递阶辨识思想并结合最小二乘和粒子群算法,建立针对有理模型的混合算法辨识理论框架;

2)所提出的智能多步长算法能够克服粒子群混合算法计算量大、计算精度不高的问题,提高模型的辨识效率;

3)利用柔性最小二乘算法,克服混合辨识算法中有理模型分母结构已知的假设条件,所提出的在线算法能够利用最新数据更新参数,因此更具有实际应用前景.

## 1 问题描述

考虑如下有理模型:

$$y(t) = \frac{\mathbf{a}(t)}{\mathbf{b}(t)} + v(t). \quad (1)$$

其中: $y(t)$ 是系统输出; $v(t)$ 是均值为零、方差为 $\sigma^2$ 的高斯白噪声; $\mathbf{a}(t)$ 和 $\mathbf{b}(t)$ 是非线性项,可以表示为

$$\mathbf{a}(t) = \psi^T(t)\theta_a, \quad \mathbf{b}(t) = \varphi^T(t)\theta_b,$$

$$\psi(t) = [\psi_1(t), \dots, \psi_m(t)]^T \in \mathbf{R}^m,$$

$$\varphi(t) = [\varphi_1(t), \dots, \varphi_n(t)]^T \in \mathbf{R}^n,$$

$$\theta_a = [a_1, \dots, a_m]^T \in \mathbf{R}^m,$$

$$\theta_b = [b_1, \dots, b_n]^T \in \mathbf{R}^n.$$

$\psi_i(t)(i = 1, \dots, m)$ 和 $\varphi_j(t)(j = 1, \dots, n)$ 是包含输入 $u(t), \dots, u(1)$ 和输出 $y(t-1), \dots, y(1)$ 的非线性项,其可由 $u^2(t)$ 、 $u(t-1)y(t-2)$ 、 $y(t-1)u(t)$ 等表示; $\theta_a$ 和 $\theta_b$ 是待辨识的参数向量.

传统的辨识算法是假设 $b_1 = 1$ ,从而将复杂的有理模型转换为线性自回归模型,即将模型(1)简化为

$$\begin{aligned} \bar{y}(t) = & \psi^T(t)\theta_a - \varphi^T(t)\theta_b y(t) + \varphi_1(t)y(t) + \mathbf{b}(t)v(t) = \\ & a_1\psi_1(t) + \dots + a_m\psi_m(t) - b_2\varphi_2(t)y(t) - \dots - \\ & b_n\varphi_n(t)y(t) + \mathbf{b}(t)v(t), \end{aligned} \quad (2)$$

其中 $\bar{y}(t) = y(t)\varphi_1(t)$ .

上述线性自回归模型可以简化为

$$\bar{y}(t) = \phi^T(t)\vartheta + \zeta(t). \quad (3)$$

其中

$$\begin{aligned} \phi(t) = & [\psi_1(t), \dots, \psi_m(t), -\varphi_2(t)y(t), \dots, \\ & -\varphi_n(t)y(t)]^T \in \mathbf{R}^{m+n-1}, \end{aligned}$$

$$\vartheta = [a_1, \dots, a_m, b_2, \dots, b_n]^T \in \mathbf{R}^{m+n-1},$$

$$\zeta(t) = \mathbf{b}(t)v(t).$$

因为 $v(t)$ 是独立同分布的高斯白噪声,且 $v(t)$ 与 $t$ 时刻之前的输入和输出是不相关的,所以

$$\mathbf{E}[\zeta(t)] = \mathbf{E}[\mathbf{b}(t)]\mathbf{E}[v(t)] = 0.$$

因为信息向量中的后  $n-1$  个向量包含  $y(t)$ , 而  $y(t)$  是与  $v(t)$  相关的, 所以信息向量  $\phi(t)$  与噪声  $\zeta(t)$  是相关的.

假设收集到  $L$  组输入和输出数据, 并定义

$$\begin{aligned}\bar{Y}(L) &= [\bar{y}(L), \dots, \bar{y}(1)]^T \in \mathbf{R}^L, \\ \Phi(L) &= [\phi(L), \dots, \phi(1)] \in \mathbf{R}^{(m+n-1) \times L}, \\ \zeta(L) &= [\zeta(L), \dots, \zeta(1)]^T \in \mathbf{R}^L.\end{aligned}$$

有理模型可以写为

$$\bar{Y}(L) = \Phi^T(L)\vartheta + \zeta(L).$$

利用最小二乘算法辨识系统参数, 得到

$$\hat{\vartheta} = [\Phi^T(L)\Phi(L)]^{-1}\Phi^T(L)\bar{Y}(L). \quad (4)$$

估计参数的期望为

$$\begin{aligned}E[\hat{\vartheta}] &= \\ E\{[\Phi^T(L)\Phi(L)]^{-1}\Phi^T(L)\bar{Y}(L)\} &= \\ E\{[\Phi^T(L)\Phi(L)]^{-1}\Phi^T(L)[\Phi^T(L)\vartheta + \zeta(L)]\} &= \\ \vartheta + E\{[\Phi^T(L)\Phi(L)]^{-1}\Phi^T(L)\zeta(L)\}, &\quad (5)\end{aligned}$$

显然

$$E\{[\Phi^T(L)\Phi(L)]^{-1}\Phi^T(L)\zeta(L)\} \neq 0.$$

因此, 利用传统的最小二乘算法对自回归有理模型进行辨识, 其辨识结果是有偏的. 为消除噪声与信息向量相关所导致的有偏性, 可以对辨识结果进行补偿, 即所谓的偏差补偿最小二乘算法<sup>[16,18,23]</sup>.

**注1** 偏差补偿最小二乘估计法可以获得有理模型的无偏估计, 但是其存在3个缺点: 1) 由于噪声  $\zeta(t) = \mathbf{b}(t)v(t)$  受  $\mathbf{b}(t)$  影响, 其参数估计尽管是无偏的, 但是渐近收敛性得不到保障; 2) 更新参数过程中包含矩阵求逆运算, 且每一次迭代需要计算所有噪声并求其方差, 导致计算量较大; 3) 每一步迭代过程中要通过计算的噪声方差对辨识参数进行补偿, 在初始阶段由于参数估计不准确, 导致噪声估计不准确, 进而影响参数辨识效果.

## 2 有理模型的混合辨识

为减少噪声对参数辨识影响以及降低计算量, 受文献[24-25]递阶思想的启发, 本文将有理模型分解为两个子模型: 分子子模型和分母子模型, 分别对这两个子模型进行辨识以获得系统参数估计.

重新将有理模型写为

$$y(t) = \frac{\psi^T(t)\theta_a}{\varphi^T(t)\theta_b} + v(t), \quad (6)$$

定义损失函数

$$J(\theta_a, \theta_b) = \frac{1}{2} \sum_{t=1}^L \left[ y(t) - \frac{\psi^T(t)\theta_a}{\varphi^T(t)\theta_b} \right]^2.$$

假设在第  $k$  次迭代过程中, 系统在第  $k-1$  次的参数  $\hat{\theta}_a^{k-1}$ 、 $\hat{\theta}_b^{k-1}$  已经获得. 分别对  $\theta_a$  和  $\theta_b$  求导, 并令导函数为零, 即

$$\begin{aligned}\frac{\partial J(\theta_a, \hat{\theta}_b^{k-1})}{\partial \theta_a} &= \\ - \sum_{t=1}^L \frac{\psi(t)}{\varphi^T(t)\hat{\theta}_b^{k-1}} \left[ y(t) - \frac{\psi^T(t)\theta_a}{\varphi^T(t)\hat{\theta}_b^{k-1}} \right] &= 0, \quad (7)\end{aligned}$$

$$\frac{\partial J(\hat{\theta}_a^k, \theta_b)}{\partial \theta_b} =$$

$$\sum_{t=1}^L \frac{\psi^T(t)\hat{\theta}_a^k \varphi(t)}{(\varphi^T(t)\theta_b)^2} \left[ y(t) - \frac{\psi^T(t)\hat{\theta}_a^k}{\varphi^T(t)\theta_b} \right] = 0. \quad (8)$$

根据等式(7)可得

$$\hat{\theta}_a^k = \sum_{t=1}^L \left[ \frac{\psi(t)\psi^T(t)}{(\varphi^T(t)\hat{\theta}_b^{k-1})^2} \right]^{-1} \sum_{t=1}^L \left[ \frac{\psi(t)y(t)}{\varphi^T(t)\hat{\theta}_b^{k-1}} \right]. \quad (9)$$

由于  $\theta_b$  在分母上, 等式(8)不能分解为参数  $\theta_b$  的解析形式, 很难求出  $\hat{\theta}_b^k$ . 对此, 下面提出两种方法来辨识参数  $\theta_b$ .

### 2.1 基于粒子群的迭代混合算法

由于分母上含有参数, 当利用导函数对分母参数求导时, 无法分解为关于未知参数的一个解析方程, 也就无法辨识出分母参数. 本小节通过粒子群算法来辨识分母参数.

在粒子群算法中, 假设未知参数向量  $\theta_b$  是一个粒子, 在每一步都假设存在  $l$  个粒子, 例如, 在第  $k-1$  步,  $l$  个粒子分别是  $\theta_b^{k-1,1}, \dots, \theta_b^{k-1,l}$ .

在第  $k-1$  步, 每一个粒子都有自己的局部最优值 ( $\theta_{b,ob}^{k-1,i}$ , ob表示局部最优(own best)), 其满足

$$\theta_{b,ob}^{k-1,i} = \arg \min_{\theta_b} (J(\hat{\theta}_a^{k-1}, \hat{\theta}_b^{k-1,i}), J(\hat{\theta}_a^{k-1}, \hat{\theta}_b^{k-2,i})).$$

所有在  $k-1$  步的粒子都存在一个全局最优值  $\theta_{b,gb}^{k-1,i}$  (gb表示全局最优(global best)), 其满足

$$\theta_{b,gb}^{k-1} = \arg \min_{\theta_b} (J(\hat{\theta}_a^{k-1}, \hat{\theta}_b^{k-1,i}), i = 1, 2, \dots, l).$$

每一个粒子都以自己的速度向第  $k$  次的值移动, 有

$$\begin{aligned}\mathbf{V}_i^k &= \mathbf{V}_i^{k-1}w + (\theta_{b,ob}^{k-1,i} - \hat{\theta}_b^{k-1,i})c_1r_1 + \\ &(\theta_{b,gb}^{k-1} - \hat{\theta}_b^{k-1,i})c_2r_2.\end{aligned}$$

其中:  $c_1$  和  $c_2$  称为加速系数, 是介于 0 与 2 之间的常数,  $c_1$  的作用是使得粒子朝该粒子的当前最优位置靠近, 而  $c_2$  是使粒子朝全局最优靠近;  $w$  是惯性因子, 且  $w \in [0.1, 0.9]$ ;  $r_1$  和  $r_2$  是两个介于  $[0, 1]$  的任意常数. 在第  $k$  步, 每个粒子获得了一个新的值, 即

$$\hat{\theta}_b^{k,i} = \hat{\theta}_b^{k-1,i} + \mathbf{V}_i^k.$$

基于粒子群的迭代混合算法可以总结如下:

$$\hat{\theta}_b^k = \theta_{b,\text{gb}}^k = \arg \min_{\theta_b} (J(\hat{\theta}_a^k, \hat{\theta}_b^{k,i}), i = 1, 2, \dots, l); \quad (10)$$

$$\theta_{b,\text{ob}}^{k,i} = \arg \min_{\theta_b} (J(\hat{\theta}_a^k, \hat{\theta}_b^{k,i}), J(\hat{\theta}_a^k, \hat{\theta}_b^{k-1,i})); \quad (11)$$

$$\hat{\theta}_b^{k,i} = \hat{\theta}_b^{k-1,i} + \mathbf{V}_i^k, i = 1, 2, \dots, l; \quad (12)$$

$$\mathbf{V}_i^k = \mathbf{V}_i^{k-1}w + (\theta_{b,\text{ob}}^{k-1,i} - \hat{\theta}_b^{k-1,i})c_1r_1 + (\theta_{b,\text{gb}}^{k-1} - \hat{\theta}_b^{k-1,i})c_2r_2. \quad (13)$$

**注2** 粒子群迭代混合算法结合递阶辨识思想解决了系统分母上存在未知参数这样的一种局限,通过最小二乘算法辨识分子参数,由粒子群算法辨识分母参数,两者交替进行,直到获得满意的参数辨识效果.同样,因信息向量中没有与噪声相关的向量,且没有改变原始噪声的统计特性,故辨识效果较好.

**注3** 粒子群迭代混合算法的辨识效果与粒子的个数密切相关,粒子个数越多,其陷入局部最优的概率越小,但是计算量越大.同样,参数 $c_1$ 、 $c_2$ 、 $w$ 、 $r_1$ 、 $r_2$ 也影响着算法的收敛速度和精度.

## 2.2 基于智能多步长搜索梯度迭代混合辨识算法

由于粒子群算法是一个智能算法,其算法的收敛性密切依赖于粒子的数目和一些不确定的常值的大小.本小节通过梯度类算法来辨识分母参数.在梯度迭代算法中,首先确定参数收敛的方向,即所谓负梯度方向,再确定该方向的步长.

假设在第 $k$ 次时,参数 $\hat{\theta}_a^k$ 已由式(9)获得,下面利用梯度迭代方法来获得参数估计 $\hat{\theta}_b^k$ .

定义损失函数

$$J(\hat{\theta}_a^k, \theta_b) = \frac{1}{2} \sum_{t=1}^L \left[ y(t) - \frac{\psi^T(t)\hat{\theta}_a^k}{\varphi^T(t)\theta_b} \right]^2. \quad (14)$$

其负梯度方向可由下式计算得到:

$$-\left. \frac{\partial J(\hat{\theta}_a^k, \theta_b)}{\partial \theta_b} \right|_{\hat{\theta}_b^{k-1}} = -\sum_{t=1}^L \frac{\psi^T(t)\hat{\theta}_a^k\varphi(t)}{(\varphi^T(t)\hat{\theta}_b^{k-1})^2} \left[ y(t) - \frac{\psi^T(t)\hat{\theta}_a^k}{\varphi^T(t)\hat{\theta}_b^{k-1}} \right].$$

因此, $\hat{\theta}_b^k$ 可以表示为

$$\hat{\theta}_b^k = \hat{\theta}_b^{k-1} - \lambda_k \sum_{t=1}^L \frac{\psi^T(t)\hat{\theta}_a^k\varphi(t)}{(\varphi^T(t)\hat{\theta}_b^{k-1})^2} \times \left[ y(t) - \frac{\psi^T(t)\hat{\theta}_a^k}{\varphi^T(t)\hat{\theta}_b^{k-1}} \right], \quad (15)$$

其中 $\lambda_k$ 是步长.步长的选择是非常重要的,如果步长太大,则导致算法发散;如果步长太小,则导致算法的收敛速度很慢.

一种通用的方法是将式(15)代入到损失函数(14)中,并求该函数关于步长 $\lambda_k$ 的导函数方程的解,

该解为当前方向的最优步长.这种方法即为经典的最速下降法.由于 $\hat{\theta}_b^k$ 在损失函数的分母上,求解关于步长的导函数方程是很困难的.因此,本文借鉴粒子群算法,提出智能多步长梯度迭代算法.即在第 $k$ 次,为步长设置一个区间, $\lambda_k \in [0, d_k]$ ,在区间内由小到大取 $l$ 个步长, $\lambda_k^1, \lambda_k^2, \dots, \lambda_k^l$ ,分别计算每个步长对应的参数值

$$\hat{\theta}_b^{k,i} = \hat{\theta}_b^{k-1} - \lambda_k^i \sum_{t=1}^L \frac{\psi^T(t)\hat{\theta}_a^k\varphi(t)}{(\varphi^T(t)\hat{\theta}_b^{k-1})^2} \times \left[ y(t) - \frac{\psi^T(t)\hat{\theta}_a^k}{\varphi^T(t)\hat{\theta}_b^{k-1}} \right], i = 1, \dots, l, \quad (16)$$

并计算其对应的损失函数的值 $J(\hat{\theta}_a^k, \hat{\theta}_b^{k,i})$ ,其中最优化步长为

$$\lambda_k^{\text{best}} = \arg \min_{\lambda_k^i} \{ J(\hat{\theta}_a^k, \hat{\theta}_b^{k,i}), i = 1, \dots, l \}. \quad (17)$$

利用智能多步长梯度迭代混合算法辨识参数的步骤如下:

- 1) 固定 $\hat{\theta}_b^{k-1}$ ,利用式(9)计算 $\hat{\theta}_a^k$ .
- 2) 在区间 $[0, d_k]$ 生成 $l$ 个步长, $\lambda_k^1 < \lambda_k^2 < \dots < \lambda_k^l$ .
- 3) 固定 $\hat{\theta}_a^k$ ,根据式(9)计算参数 $\hat{\theta}_b^{k,i}, i = 1, \dots, l$ .
- 4) 计算损失函数 $J(\hat{\theta}_a^k, \hat{\theta}_b^{k,i}), i = 1, \dots, l$ ,并找出 $\lambda_k^{\text{best}}$ .
- 5) 得到 $\hat{\theta}_b^{k,\text{best}}$ ,并令 $\hat{\theta}_b^k = \hat{\theta}_b^{k,\text{best}}$ .
- 6) 比较 $[\hat{\theta}_a^k; \hat{\theta}_b^k]$ 与 $[\hat{\theta}_a^{k-1}; \hat{\theta}_b^{k-1}]$ ,若 $\|[\hat{\theta}_a^k; \hat{\theta}_b^k] - [\hat{\theta}_a^{k-1}; \hat{\theta}_b^{k-1}]\|$ 足够小,则停止运算,获得估计值 $[\hat{\theta}_a^k; \hat{\theta}_b^k]$ ;否则,令 $k = k + 1$ ,并转步骤1).

**注4** 通过将模型分解为两个子模型,并分别交互估计每一部分参数:分子通过最小二乘算法获得参数,分母通过智能多步长梯度迭代算法计算参数.因此,相对于将模型转换成线性自回归模型的传统辨识方法,智能多步长梯度迭代混合算法计算量小,且是无偏估计.

**注5** 利用智能多步长梯度方法过程中,影响算法收敛性和收敛速度主要有两个因素:1)步长区间 $[0, d_k]$ ;2)步长个数 $l$ .

1) 如果 $d_k$ 选择过大,则由最小的步长 $\lambda_k^1$ 计算的损失函数可能满足 $J(\hat{\theta}_a^k, \hat{\theta}_b^{k,1}) > J(\hat{\theta}_a^{k-1}, \hat{\theta}_b^{k-1})$ ,即算法发散;

2) 如果 $d_k$ 选择过小,所有步长计算出的损失函数 $J(\hat{\theta}_a^k, \hat{\theta}_b^{k,i}) < J(\hat{\theta}_a^{k-1}, \hat{\theta}_b^{k-1}), i = 1, \dots, l$ ,此时算法是收敛的,但是其收敛速度较慢.

针对上述问题,若步长选择过大,则令新的步长上限 $d_k^{\text{new}} = \frac{1}{2}d_k$ ;若选择过小,则 $d_k^{\text{new}} = 2d_k$ .在实际应用中,当参数越来越靠近真值时,步长将越来越小,

因此可以逐渐减小  $d_k$ . 即在初始阶段选择一个大的常值  $d$ , 使  $d_1 = d$ , 在第  $k$  次时,  $d_k = \frac{k}{1+2+\dots+k}d$ .

**注6** 相比较于粒子群迭代混合算法, 智能多步长算法计算量小, 算法稳定, 但是存在易陷入局部最优的问题, 即一旦初始值选择不当, 则辨识值容易陷入局部最优. 而粒子群迭代混合算法可以通过选择大量的初始粒子, 进而降低估计值陷入局部最优的概率.

### 2.3 混合辨识算法收敛性分析

混合辨识算法中, 假设第  $k-1$  次的参数辨识结果为  $\hat{\theta}_a^{k-1}$  和  $\hat{\theta}_b^{k-1}$ , 则其对应的损失函数为

$$J(\hat{\theta}_a^{k-1}, \hat{\theta}_b^{k-1}) = \frac{1}{2} \sum_{t=1}^L \left[ y(t) - \frac{\psi^T(t)\hat{\theta}_a^{k-1}}{\varphi^T(t)\hat{\theta}_b^{k-1}} \right]^2.$$

固定  $\hat{\theta}_b^{k-1}$ , 则关于未知向量  $\theta_a$  的损失函数为

$$J(\theta_a, \hat{\theta}_b^{k-1}) = \frac{1}{2} \sum_{t=1}^L \left[ y(t) - \frac{\psi^T(t)\theta_a}{\varphi^T(t)\hat{\theta}_b^{k-1}} \right]^2.$$

在粒子群迭代混合算法以及智能多步长搜索梯度迭代混合辨识算法中, 更新分子参数时是利用最小二乘算法, 即

$$\hat{\theta}_a^k = \arg \min_{\theta_a} [J(\theta_a, \hat{\theta}_b^{k-1})],$$

因此可得

$$J(\hat{\theta}_a^k, \hat{\theta}_b^{k-1}) \leq J(\hat{\theta}_a^{k-1}, \hat{\theta}_b^{k-1}). \quad (18)$$

固定  $\hat{\theta}_a^k$ , 则关于未知参数向量  $\theta_b$  的损失函数为

$$J(\hat{\theta}_a^k, \theta_b) = \frac{1}{2} \sum_{t=1}^L \left[ y(t) - \frac{\psi^T(t)\hat{\theta}_a^k}{\varphi^T(t)\theta_b} \right]^2.$$

而粒子群算法和智能多步长算法中都希望找到合适的粒子或步长使得

$$J(\hat{\theta}_a^k, \hat{\theta}_b^k) \leq J(\hat{\theta}_a^k, \hat{\theta}_b^{k-1}). \quad (19)$$

根据式(18)和(19)可得

$$J(\hat{\theta}_a^k, \hat{\theta}_b^k) \leq J(\hat{\theta}_a^{k-1}, \hat{\theta}_b^{k-1}).$$

即随着迭代次数  $k$  的增加, 损失函数  $J(\theta_a, \theta_b)$  是单调递减的, 因此, 混合辨识算法是收敛的.

**注7** 更新分子参数时是利用最小二乘算法, 因此更新得到的参数向量  $\hat{\theta}_a^k$  满足  $\hat{\theta}_a^k = \arg \min_{\theta_a} [J(\theta_a, \hat{\theta}_b^{k-1})]$ ; 而在更新分母参数过程中, 只需找到合适的  $\hat{\theta}_b^k$ , 使得  $J(\hat{\theta}_a^k, \hat{\theta}_b^k) < J(\hat{\theta}_a^k, \hat{\theta}_b^{k-1})$  即可, 如果找不到, 则令  $\hat{\theta}_b^k = \hat{\theta}_b^{k-1}$ , 因此, 总能使得  $J(\hat{\theta}_a^k, \hat{\theta}_b^k) \leq J(\hat{\theta}_a^{k-1}, \hat{\theta}_b^{k-1})$ .

**注8** 如果损失函数存在唯一的极小值点 ( $\theta_a^{\text{true}}, \theta_b^{\text{true}}$ ), 由于损失函数是单调递减的, 则根据压缩映射原理, 混合辨识方法更新得到的参数值会渐近收敛到最小值点; 如果损失函数存在多个极小值点, 则混合辨识方法可能会收敛到局部最优. 此时, 可以通过选

择更多的粒子或多个初始值来使得估计值尽可能收敛到最优值.

### 3 有理模型的柔性最小二乘算法

混合辨识算法借助递阶辨识思想分别辨识分子和分母参数. 很显然, 分子的辨识效果影响分母的辨识精度和速度, 同样, 分母的辨识结果也影响着分子. 另外, 混合辨识方法要求分母模型结构已知, 因此限制了其使用范围. 本节借助柔性最小二乘方法来消除混合辨识方法的缺点.

将有理模型(1)整理为

$$\begin{aligned} y(t) &= \frac{\mathbf{a}(t)}{\mathbf{b}(t)} + v(t) = \\ &= \frac{a_1}{\mathbf{b}(t)}\psi_1(t) + \dots + \frac{a_m}{\mathbf{b}(t)}\psi_m(t) + v(t) = \\ &= a_1(t)\psi_1(t) + \dots + a_m(t)\psi_m(t) + v(t), \end{aligned} \quad (20)$$

其中  $a_i(t) = \frac{a_i}{\mathbf{b}(t)}$  ( $i = 1, \dots, m$ ) 是时变参数. 其辨识的目的是通过收集的输入和输出数据把系统的可变参数辨识出来.

将系统转化为

$$\begin{aligned} y(t) &= \psi^T(t)\theta(t) + v(t), \\ \psi(t) &= [\psi_1(t), \dots, \psi_m(t)]^T, \\ \theta(t) &= [a_1(t), \dots, a_m(t)]^T. \end{aligned}$$

假设一共收集了  $L$  组输入和输出数据, 则存在  $L$  组待辨识的参数向量  $\theta(1), \dots, \theta(L)$ . 定义损失函数

$$\begin{aligned} J(\theta(L), \theta(L-1), \dots, \theta(1)) &= \\ &= \frac{1}{2} \sum_{t=1}^L [y(t) - \psi^T(t)\theta(t)]^2 + \\ &+ \frac{\mu}{2} \sum_{t=1}^{L-1} [\theta(t+1) - \theta(t)]^T [\theta(t+1) - \theta(t)], \end{aligned} \quad (21)$$

其中  $\mu$  是权值系数, 受参数变化速度的影响. 将损失函数分别对每一组参数向量求偏导, 并利用导函数为零求得每一组参变量的估计值.

为一次性辨识出系统的参数, 定义下列参数向量、信息矩阵以及噪声向量:

$$\Theta(L) = [\theta^T(L), \dots, \theta^T(1)]^T \in \mathbf{R}^{Lm \times 1},$$

$$\Psi(L) = \begin{bmatrix} \psi(L) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \psi(1) \end{bmatrix} \in \mathbf{R}^{Lm \times L},$$

$$V(L) = [v(L), \dots, v(1)]^T.$$

于是有理模型可进一步转化为

$$Y(L) = \Psi^T(L)\Theta(L) + V(L).$$

定义一个扩展矩阵

$$R_t(\mu) = \begin{cases} \psi(1)\psi^T(1) + \mu I, & t = 1; \\ \psi(t)\psi^T(t) + 2\mu I, & t \neq 1, L; \\ \psi(L)\psi^T(L) + \mu I, & t = L. \end{cases} \quad (22)$$

$$R(L, \mu) = \begin{bmatrix} R_L(\mu) & -\mu I & 0 & \dots & 0 \\ -\mu I & R_{L-1}(\mu) & -\mu I & \dots & 0 \\ 0 & -\mu I & R_{L-2}(\mu) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & -\mu I \\ 0 & 0 & \dots & -\mu I & R_1(\mu) \end{bmatrix} \in \mathbf{R}^{P \times P},$$

$$P = Lm. \quad (23)$$

损失函数(21)可以转化为

$$J(\Theta(L)) = \frac{1}{2} \Theta^T(L) R(L, \mu) \Theta(L) - \Theta^T(L) \Psi(L) Y(L) + \frac{1}{2} Y^T(L) Y(L).$$

因此,参数向量 $\Theta(L)$ 可由下式计算:

$$\Theta(L) = R^{-1}(L, \mu) \Psi(L) Y(L). \quad (24)$$

**注9** 当权值 $\mu > 0$ 时,矩阵 $R(L, \mu)$ 是对称正定矩阵,此矩阵的逆始终存在. 因此,由式(24)可以求得参数的估计值.

**注10** 式(24)是通过离线方法来估计系统参数的,必须要求 $L \geq m$ ,进而求解矩阵的逆. 要使参数估计精度提高, $L$ 必须满足 $L \gg m^{[26-27]}$ ,由此带来一个新的问题,矩阵 $R(L, \mu)$ 的阶数会急剧增长,导致计算量非常大. 例如,当系统存在 $m = 10$ 个参数,收集了 $L = 100$ 组数据时,矩阵 $R(L, \mu)$ 的阶数为1000. 此时,如果只统计乘法和除法,则求解此矩阵的逆大致需要 $1000^3 = 10^9$ 次乘法和除法运算.

为减小计算量,下面提出在线算法将系统参数辨识出来. 假设在时间 $t$ ,已采集到当前时刻的输入 $u(t)$ 和输出 $y(t)$ ,并且之前所有时刻的参数已知,即 $\hat{\theta}(1), \dots, \hat{\theta}(t-1)$ 已知,于是损失函数为

$$J(\hat{\theta}(1), \dots, \hat{\theta}(t-1), \theta(t)) = \frac{1}{2} \sum_{i=1}^{t-1} [y(i) - \psi^T(i)\hat{\theta}(i)]^2 + \frac{1}{2} [y(t) - \psi^T(t)\theta(t)]^2 + \frac{\mu}{2} \sum_{i=1}^{t-1} [\hat{\theta}(i) - \hat{\theta}(i-1)]^T [\hat{\theta}(i) - \hat{\theta}(i-1)] + \frac{\mu}{2} [\theta(t) - \hat{\theta}(t-1)]^T [\theta(t) - \hat{\theta}(t-1)]. \quad (25)$$

由于 $t$ 时刻之前的所有参数向量已知,根据

$$\frac{\partial J(\hat{\theta}(1), \dots, \hat{\theta}(t-1), \theta(t))}{\partial \theta(t)} = 0,$$

可得

$$\hat{\theta}(t) = [\psi(t)\psi^T(t) + \mu I]^{-1} [\psi(t)y(t) + \mu \hat{\theta}(t-1)] = \hat{\theta}(t-1) + [\psi(t)\psi^T(t) + \mu I]^{-1} \psi(t) \times [y(t) - \psi(t)\psi^T(t)\hat{\theta}(t-1)]. \quad (26)$$

利用滤波柔性递推最小二乘算法辨识参数 $\theta(t)$ 的步骤如下:

- 1)初始化:假设 $u(t) = 0, y(t) = 0, t \leq 0, \hat{\theta}(0) = 1/10^6$ ,其中 $\mathbf{1} = [1, 1, \dots, 1]^T \in \mathbf{R}^m$ .
- 2)选择一个合适的正的常数 $\mu$ ,并令 $t = 1$ .
- 3)采集输入和输出数据 $u(t)$ 和 $y(t)$ .
- 4)构建损失函数 $J(\hat{\theta}(1), \dots, \hat{\theta}(t-1), \theta(t))$ .
- 5)由式(26)更新参数得到 $\hat{\theta}(t)$ .
- 6)若 $t < L$ ,则转步骤3);否则,得到估计出的一组参数向量 $\hat{\theta}(1), \dots, \hat{\theta}(L)$ .

对于权值 $\mu$ 的选择,有以下3种情形:

1)若 $\mu = 0$ ,则矩阵 $[\psi(t)\psi^T(t)]$ 是奇异的,此时根据式(26)无法求解出参数值. 同样,由损失函数(25)也可知,导函数的方程组未知数个数多于方程个数,此时无法求解方程的解(参数值).

2)若 $\mu \neq 0$ ,则:当系统的参数变化较快时,应赋较小的权值给参数误差,即此时要使 $\mu$ 取尽量小的正常数;当参数变化较慢时,应赋较大的权值 $\mu$ .

3)当 $\mu \rightarrow \infty$ 时,测量误差函数几乎对参数辨识不起作用,此时参数辨识值始终保持不变. 同样,根据式(26)可知, $\hat{\theta}(L) = \hat{\theta}(L-1) = \dots = \hat{\theta}(1) = \hat{\theta}(0)$ .

**注11** 柔性递推最小二乘算法适合于系统参数变化较慢的情形,即在本文中系统的分母 $b(t)$ 缓慢变化的情形. 若系统的分母变化较快,且分母结构已知,则利用柔性最小二乘算法其辨识效果得不到保障,此时可以用混合算法来辨识系统的参数(如果损失函数只存在唯一的极值点,则优先考虑智能多步长梯度迭代混合辨识算法;否则,考虑使用粒子群迭代混合算法). 因此,可以根据不同的情形选择不同的算法.

**注12** 柔性递推最小二乘算法待辨识的参数维数小于基于递阶辨识思想的混合迭代法待辨识参数维数,且是一个在线辨识算法,因此,柔性递推最小二乘算法计算量小于混合迭代算法的计算量.

## 4 仿真例子

### 4.1 混合辨识算法

考虑如下有理模型:

$$y(t) = \frac{0.3y(t-1) + 0.7u(t-1)}{1 + u(t-1)} + v(t),$$

$$\theta = [a_1, a_2, b_1, b_2]^T = [0.3, 0.7, 1, 1]^T,$$

$$\vartheta = [a_1, a_2, b_2]^T = [0.3, 0.7, 1]^T.$$

输入  $\{u(t)\}$  满足均值为零、方差为1的持续激励信号.  $\{v(t)\}$  是均值为零、方差为  $\sigma^2 = 0.10^2$  的白噪声信号, 仿真时共采集  $L = 100$  组输入和输出数据. 采用传统最小二乘迭代法 (least squares iterative algorithm, LS) 和文献 [16] 所提出的偏差补偿最小二乘迭代法 (bias compensation least squares iterative algorithm, BCLS), 参数估计值和估计误差  $\delta = \|\hat{\vartheta}^k - \vartheta\|/\|\vartheta\|$  随迭代次数  $k$  变化的结果如表1所示.

表1 最小二乘和偏差补偿最小二乘参数估计

方法	$k$	$a_1$	$a_2$	$b_2$	$\delta/\%$
LS	2	0.33093	0.63447	0.89576	10.09941
	10	0.33093	0.63447	0.89576	10.09941
	20	0.33093	0.63447	0.89576	10.09941
	30	0.33093	0.63447	0.89576	10.09941
BCLS	2	0.32298	0.65501	0.93761	6.95212
	10	0.30131	0.67892	0.96646	3.15300
	20	0.30131	0.67892	0.96646	3.15300
	30	0.30131	0.67892	0.96646	3.15300
true values		0.30000	0.70000	1.00000	

1) 利用粒子群迭代混合算法 (particle swarm optimization compound, PSO-C) 辨识有理模型参数 [(PSO-C-1,  $l = 50$ ), (PSO-C-2,  $l = 100$ )].  $c_1 = 1.6$ ,  $c_2 = 1.4962$ ,  $w = 0.7298$ ,  $r_1(k) = r_2(k) = \text{rand}/k$ . 参数估计值、真实值以及它们之间的误差  $\delta = \|\hat{\theta}_a^k; \hat{\theta}_b^k\| - [\theta_a; \theta_b]\|/\|[\theta_a; \theta_b]\|$  随迭代次数变化的结果如表2所示. 表2表明, 粒子群迭代混合算法能分别辨识出有理模型分子和分母的参数, 且粒子数越多, 辨识越精确, 但其相应的计算量也越大.

表2 粒子群迭代混合算法参数估计

算法	$k$	$a_1$	$a_2$	$b_1$	$b_2$	$\delta/\%$
PSO-C-1	10	0.15879	1.02057	1.12860	1.65653	47.01448
	20	0.42396	0.44767	1.12860	1.65653	45.17861
	30	0.29273	0.71368	0.96493	1.07619	5.31034
	40	0.29926	0.69973	0.97406	1.02975	2.45779
PSO-C-2	10	0.31829	0.76153	0.99691	0.98986	4.05031
	20	0.27602	0.61164	0.99691	0.98986	5.73799
	30	0.28645	0.64725	0.99691	0.98986	3.45437
	40	0.29324	0.67086	0.99691	0.98986	1.97595
true values		0.30000	0.70000	1.00000	1.00000	

2) 利用智能多步长梯度迭代混合辨识算法 (intelligent multi-step-length gradient iterative, I-M-GI) 辨识有理模型 [(I-M-GI-1,  $d = 0.05$ ,  $l = 50$ ), (I-M-GI-2,  $d = 0.15$ ,  $l = 10$ ), (I-M-GI-3,  $d = 0.15$ ,  $l = 50$ )]. 参数估计值、真实值以及它们之间的误差  $\delta = \|\hat{\theta}^k - \vartheta\|/\|\vartheta\|$  随迭代次数变化的结果如表3所示. 由表3可知, 步长区间的大小和步长的个数影响

着辨识效果, 大的区间上界以及足够多的步长数会加快参数收敛的速度.

表3 智能多步长梯度迭代混合辨识算法参数估计

算法	$k$	$a_1$	$a_2$	$b_1$	$b_2$	$\delta/\%$
I-M-GI-1	2	0.35658	0.99753	1.95209	1.95519	86.05474
	10	0.34885	0.94518	1.75377	1.76957	68.84747
	20	0.33761	0.87739	1.50913	1.53970	47.55119
	30	0.32553	0.81267	1.29065	1.33166	28.38156
I-M-GI-2	2	0.35479	0.98486	1.87254	1.88078	79.27129
	10	0.33278	0.85061	1.37804	1.41577	36.27732
	20	0.30929	0.73562	1.05702	1.09432	7.23450
	30	0.30202	0.70399	1.00658	1.01339	0.96985
I-M-GI-3	2	0.35477	0.98475	1.85587	1.86519	77.88762
	10	0.32784	0.82451	1.28712	1.32890	28.31781
	20	0.30446	0.71443	1.01587	1.03825	2.74411
	30	0.30065	0.69823	1.00331	0.99709	0.29869
true values		0.30000	0.70000	1.00000	1.00000	

表2和表3表明: 智能多步长梯度迭代混合辨识算法更加稳定和精确, 这是因为该算法方向是最优的, 所以每次只需合适的步长就能保证算法是渐近收敛的; 而粒子群迭代混合算法影响其收敛的因素很多, 具有不确定性. 因此, 为保证粒子群算法的稳定性, 需要充分多的粒子数.

表1~表3说明: 1) 偏差补偿最小二乘迭代法和混合迭代算法能获得较好的辨识效果, 而最小二乘迭代法由于没有考虑偏差补偿, 辨识效果是最差的, 特别是当噪声方差较大时, 其辨识效果会急剧恶化; 2) 混合迭代法对分母第1项的参数  $b_1$  无限制条件 (其可以辨识出  $b_1$  的值), 而传统最小二乘或偏差最小二乘要假设分母第1项参数  $b_1 = 1$ .

## 4.2 柔性最小二乘辨识算法

将4.1节有理模型转化为时变参数模型

$$y(t) = a_1(t)y(t-1) + a_2(t)u(t-1) + v(t),$$

$$\theta(t) = [a_1(t), a_2(t)]^T = \begin{bmatrix} 0.3 & 0.7 \\ \mathbf{b}(t) & \mathbf{b}(t) \end{bmatrix}^T,$$

$$\mathbf{b}(t) = 1 + u^2(t-1),$$

$$\psi(t) = [y(t-1), u(t-1)]^T.$$

输入采用  $u(t) = \sin\left(\frac{\pi}{20}t\right)$ ,  $\{v(t)\}$  是均值为零、方差为  $\sigma^2 = 0.10^2$  的白噪声信号. 仿真时共采集  $L = 100$  组数据, 相应地, 100组参数  $\theta(1), \dots, \theta(100)$  需要辨识.

利用滤波柔性递推最小二乘算法 (filtering flexible recursive least squares, F-F-RLS) 辨识系统参数 ( $\mu = 0.1$ ). 参数估计值、输出估计值以及它们之间相应的误差分别如图1、图2所示. 图1表明, 当系统参数缓慢变化时, 滤波柔性递推最小二乘算法的参数估

计值能很好地跟踪到真实的参数值. 图2表明, 估计输出能逼近真实输出.

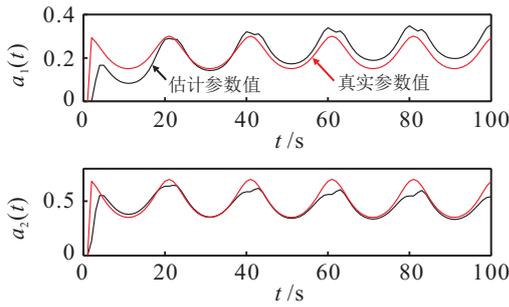


图1 参数估计随t变化曲线(F-F-RLS)(1)

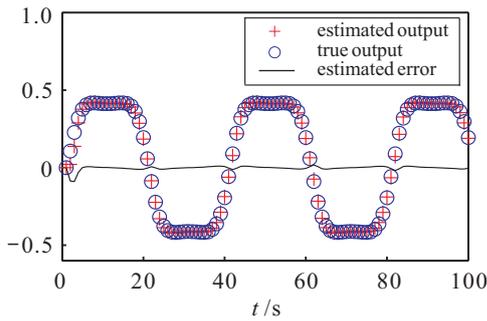


图2 输出估计、真实输出及其误差变化曲线(1)

进一步, 利用蒙特卡罗仿真实验分别验证3种算法对噪声的鲁棒性(20组噪声), 仿真结果如图3~图5所示. 图3~图5表明, 3种算法对噪声都具有较强的鲁棒性. 3种算法蒙特卡罗仿真耗时如表4所示(计算机配置Intel(R) Core(TM) i5-7 220U: 2.50 GHz, 2.71 GHz; RAM: 8.0 GB; Windows 10). 表4表明, F-F-RLS算法计算量最小, 耗时最短; 其次是I-M-GI-3算法; 计算量最大、耗时最多的是PSO-C-2算法.

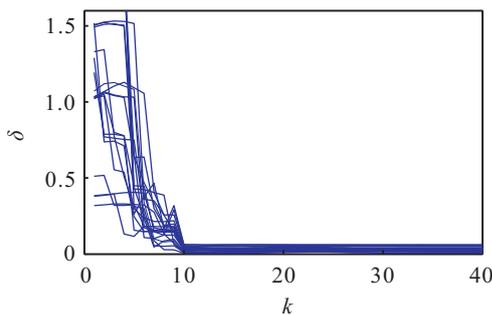


图3 参数误差 $\delta$ 随k变化曲线(PSO-C-2)

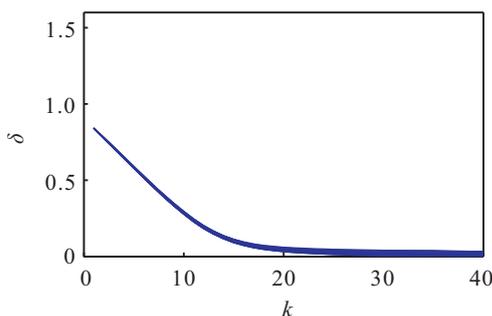


图4 参数误差 $\delta$ 随k变化曲线(I-M-GI-3)

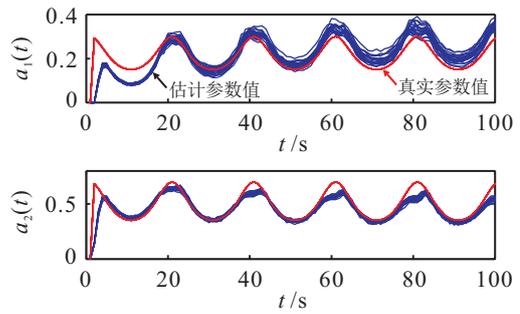


图5 参数估计随t变化曲线(F-F-RLS)(2)

表4 3种算法仿真耗时

算法	PSO-C-2	I-M-GI-3	F-F-RLS
t/s	23.520	20.331	1.288

### 5 结论

针对复杂的有理模型, 本文提出两类辨识方法. 一类是基于递阶辨识思想的混合迭代法, 将模型分解为分子模型和分母模型, 对分子模型利用最小二乘法迭代法辨识其参数, 对分母模型分别使用粒子群搜索算法和智能多步长梯度迭代算法, 进而分子和分母交互辨识. 由于将模型分解成两个子模型, 计算量大减少. 进一步, 为消除模型结构已知的假设, 充分利用新的数据在线更新参数, 又提出了基于柔性递推最小二乘的辨识算法. 把有理模型视为时变参数线性系统, 构造关于输出误差和参数误差的损失函数, 并为每一类误差分配不同的权值: 当参数缓慢变化时, 参数误差分配较大的权值; 反之, 分配较小的权值. 该算法在系统参数缓慢变化时, 具有良好的辨识效果.

本文提出的有理模型辨识的两种方法, 为实现有理模型的精确控制和故障诊断提供了理论模型基础. 进一步可以将该方法推广到复杂非线性系统以及切换系统辨识中, 对不断涌现的工程应用提供必要的理论模型支撑.

### 参考文献(References)

- [1] Billings S A, Zhu Q M. Rational model identification using an extended least-squares algorithm[J]. International Journal of Control, 1991, 54(3): 529-546.
- [2] Mu B Q, Bai E W, Zheng W X, et al. A globally consistent nonlinear least squares estimator for identification of nonlinear rational systems[J]. Automatica, 2017, 77: 322-335.
- [3] Geng X P, Zhu Q M, Liu T, et al. U-model based predictive control for nonlinear processes with input delay[J]. Journal of Process Control, 2019, 75: 156-170.
- [4] Klipp E, Herwig R, Kowald A, et al. Systems biology in practice: Concepts, implementation and application[M]. Weinheim: Wiley-VCH, 2005.
- [5] 朱全民. 非线性系统辨识[J]. 控制理论与应用, 1994,

- 11(6): 641-652.  
(Zhu Q M. Nonlinear system identification[J]. Control Theory & Applications, 1994, 11(6): 641-652.)
- [6] Söderström T, Soverini U. Errors-in-variables identification using an maximum likelihood estimation in the frequency domain[J]. Automatica, 2017, 79: 131-143.
- [7] 王宏伟, 王佳, 夏浩. 基于子空间方法的非均匀周期刷新和采样系统辨识[J]. 控制与决策, 2014, 29(5): 901-906.  
(Wang H W, Wang J, Xia H. Identification of non-uniform period refresh and sampling system via subspace method[J]. Control and Decision, 2014, 29(5): 901-906.)
- [8] Zhu Q M, Wang Y J, Zhao D Y, et al. Review of rational (total) nonlinear dynamic system modelling, identification, and control[J]. International Journal of Systems Science, 2015, 46(12): 2122-2133.
- [9] 丁锋, 汪菲菲. 损失数据线性参数系统的递推最小二乘辨识方法[J]. 控制与决策, 2016, 31(12): 2261-2266.  
(Ding F, Wang F F. Recursive least squares identification algorithms for linear-in-parameter systems with missing data[J]. Control and Decision, 2016, 31(12): 2261-2266.)
- [10] Chen G Y, Gan M, Chen C L P, et al. A regularized variable projection algorithm for separable nonlinear least-squares problems[J]. IEEE Transactions on Automatic Control, 2019, 64(2): 526-537.
- [11] 丁锋. 系统辨识算法的复杂性、收敛性及计算效率研究[J]. 控制与决策, 2016, 31(10): 1729-1741.  
(Ding F. Complexity, convergence and computational efficiency for system identification algorithms[J]. Control and Decision, 2016, 31(10): 1729-1741.)
- [12] 徐玲. 基于移动数据窗的传递函数多新息随机梯度辨识方法[J]. 控制与决策, 2017, 32(6): 1091-1096.  
(Xu L. Moving data window based multi-innovation identification stochastic gradient parameter identification method for transfer functions[J]. Control and Decision, 2017, 32(6): 1091-1096.)
- [13] 衷路生, 李兵, 龚锦红, 等. 高速列车非线性模型的极大似然辨识[J]. 自动化学报, 2014, 40(12): 2950-2958.  
(Zhong L S, Li B, Gong J H, et al. Maximum likelihood identification of nonlinear models for high-speed train[J]. Acta Automatica Sinica, 2014, 40(12): 2950-2958.)
- [14] 王呈, 陈晶, 荀径, 等. 基于混合滤波最大期望算法的高速列车建模[J]. 自动化学报, 2019, 45(12): 2260-2267.  
(Wang C, Chen J, Xun J, et al. Hybrid filter based expectation maximization algorithm for high-speed train modeling[J]. Acta Automatica Sinica, 2019, 45(12): 2260-2267.)
- [15] Chen J, Huang B, Ding F, et al. Variational Bayesian approach for ARX systems with missing observations and varying time-delays[J]. Automatica, 2018, 94: 194-204.
- [16] Billings S A, Zhu Q M. A structure detection algorithm for nonlinear dynamic rational models[J]. International Journal of Control, 1994, 59(6): 1439-1463.
- [17] Zhu Q M. A back propagation algorithm to estimate the parameters of non-linear dynamic rational models[J]. Applied Mathematical Modelling, 2003, 27(3): 169-187.
- [18] Chen J, Zhu Q M, Li J, et al. Biased compensation recursive least squares-based threshold algorithm for time-delay rational models via redundant rule[J]. Nonlinear Dynamics, 2018, 91(2): 797-807.
- [19] Chen J, Ding F, Zhu Q M, et al. Maximum likelihood based identification methods for rational models[J]. International Journal of Systems Science, 2019, 50(14): 2579-2591.
- [20] Kalaba R, Tesfatsion L. Time-varying linear regression via flexible least squares[J]. Computers & Mathematics with Applications, 1989, 17(8): 1215-1245.
- [21] Kalaba R, Tesfatsion L. A multicriteria approach to model specification and estimation[J]. Computational Statistics & Data Analysis, 1996, 21(2): 193-214.
- [22] Montana G, Triantafyllopoulos K, Tzagaris T. Flexible least squares for temporal data mining and statistical arbitrage[J]. Expert Systems with Applications, 2009, 36(2): 2819-2830.
- [23] 丁锋. 系统辨识新论[M]. 北京: 科学出版社, 2013.  
(Ding F. System identification—New theory and methods[M]. Beijing: Science Press, 2013.)
- [24] Xu L, Ding F, Zhu Q M. Hierarchical Newton and least squares iterative estimation algorithm for dynamic systems by transfer functions based on the impulse responses[J]. International Journal of Systems Science, 2019, 50(1): 141-151.
- [25] 代冀阳, 王村松, 殷林飞, 等. 飞行器分层势场路径规划算法[J]. 控制理论与应用, 2015, 32(11): 1505-1510.  
(Dai J Y, Wang C S, Yin L F, et al. Hierarchical potential field algorithm of path planning for aircraft[J]. Control Theory & Applications, 2015, 32(11): 1505-1510.)
- [26] Chen H F, Guo L. Identification and stochastic adaptive control[M]. Boston: Birkhauser, 1991.
- [27] 丁锋. 系统辨识——辨识方法性能分析[M]. 北京: 科学出版社, 2014: 55-61.  
(Ding F. System identification—Performance analysis for identification methods[M]. Beijing: Science Press, 2014: 55-61.)

## 作者简介

陈晶(1981—), 男, 教授, 博士, 从事系统辨识、机器学习等研究, E-mail: chenjing1981929@126.com;

朱全民(1955—), 男, 教授, 博士生导师, 从事有理模型的辨识与控制等研究, E-mail: quan.zhu@uwe.ac.uk.

(责任编辑: 李君玲)