

基于厚尾双学生氏t分布的非线性状态空间系统 鲁棒辨识方法

刘鑫¹, 海洋², 代伟^{1,2}

(1. 中国矿业大学人工智能研究院, 江苏徐州 221116; 2. 中国矿业大学信息与控制工程学院, 江苏徐州 221116)

摘要: 状态空间模型作为一种常见且重要的模型结构在自动化领域有着广泛的应用, 本文针对异常值干扰下的非线性状态空间系统辨识问题开展研究. 与现有的辨识方法不同, 本文充分考虑了状态转移过程和输出量测过程均受到异常值干扰的情况, 提出了一种更加全面的鲁棒辨识算法. 首先利用两个相互独立的学生氏t分布分别对状态噪声和输出噪声进行建模以保障算法的鲁棒性; 其次利用粒子平滑算法估计状态变量的后验概率分布以解决状态未知问题; 最后利用期望最大化算法实现未知参数估计. 在算法实现过程中使用了学生氏t分布表达式的数学分解, 这样做的好处是: (1) 更加有利于算法的推导和实现; (2) 更清晰地解释了算法的鲁棒性能. 并且本文通过数值算例和应用算例验证了该方法的有效性.

关键词: 非线性状态空间系统; 鲁棒辨识; 学生氏t分布; 粒子平滑; 期望最大化算法

基金项目: 国家自然科学基金(No.62103134, No.62373361); 国家重点研发计划(No.2022YFB3304700); 中国博士后基金(No.2023M743776)

中图分类号: TP27

文献标识码: A

文章编号: 0372-2112(2024)09-3052-13

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20230957

Robust Identification of Nonlinear State-Space System Based on Dual Heavy-Tailed Noise Distributions

LIU Xin¹, HAI Yang², DAI Wei^{1,2}

(1. Artificial Intelligence Research Institute, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China;

2. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China)

Abstract: The state space model is a common and important model structure for automation and control. In this paper, the robust identification of nonlinear state-space model corrupted by outliers is investigated. The outliers imposed on both the state transition process and the output measurement process are considered and a more comprehensive and robust identification algorithm is proposed. To ensure the robustness of the proposed algorithm, two independent heavy-tailed Student's t-distributions are used to describe the state noise and the output noise, respectively. Then the particle smoothing method is applied to estimate the posterior distribution of the unknown states. Finally, the expectation maximization algorithm is used to realize the parameter estimation problem. The mathematical decomposition of the Student's t-distribution is employed in the identification process which brings two main advantages: (1) facilitating the derivation and implementation of the proposed algorithm; (2) providing a more clearer explanation of the robustness of the algorithm. The usefulness of the proposed algorithm is demonstrated via the numerical and mechanical examples.

Key words: nonlinear state-space model; robust system identification; Student's t-distribution; particle smoother; expectation maximization algorithm

Foundation Item(s): National Natural Science Foundation of China (No.62103134, No.62373361); National Key Research and Development Program of China (No. 2022YFB3304700); China Postdoctoral Science Foundation (No.2023M743776)

1 引言

系统辨识作为一类由过程数据驱动的建模方法,其优势在于不需要了解系统复杂的内部机理,而是通过采集所测系统的输入输出数据,根据一个优化准则确定所测系统等价模型^[1]. 由于绝大多数的实际工业过程都是非线性的,因此如何建立准确稳定的非线性系统模型一直是系统辨识领域的研究重点^[2-4].

系统辨识中常用的非线性模型结构包括块结构级联模型(Hammerstein、Wiener)、非线性滑动自回归模型、非线性状态空间模型等^[4-7]. 其中非线性状态空间模型作为一类描述非线性过程的常见模型,不仅能够很好地揭示系统的输入、输出变量与状态变量的联系,有利于系统控制器设计,而且状态空间模型中的状态变量也能够很好地去表征系统的内部动态^[8-11]. 因此,基于非线性状态空间模型的建模过程被广泛地应用于机器人^[12,13]、工业车轮^[14]、机电系统^[15]、连续搅拌釜式反应器^[16]等实际工业系统. 文献[16]基于高斯假设提出了一种基于缺失数据的非线性状态空间系统辨识算法,并将算法应用于连续搅拌釜式反应过程建模中. 文献[15]考虑了输出数据受到非对称噪声干扰的问题,提出了一种鲁棒辨识算法并将算法应用于机电定位系统中. 但在该研究中,状态过程仍假设服从高斯分布.

针对非线性状态空间模型的辨识问题,其主要难点在于模型的状态变量是未知的. 因此在实现模型的辨识算法之前,必须考虑估计其未知的状态变量. 目前已经提出了许多非线性状态估计方法,代表性算法包括扩展卡尔曼滤波、无迹卡尔曼滤波和容积卡尔曼滤波等^[17-22]. 以上算法大多是在高斯假设下设计实现的,但在实际生产过程中,由于现场测量设备误差、数据传输以及数据管理错误等原因,非常容易使得过程数据中含有异常离群点^[23]. 这不但降低了过程数据的质量,还会导致基于高斯假设的方法精度下降甚至算法无法收敛.

近年来,数据中的异常值问题和非高斯噪声问题受到越来越多研究人员的关注. 文献[24]通过使用最小误差熵准则代替最小均方误差准则,进而提出了最小误差熵扩展卡尔曼滤波,这提高了扩展卡尔曼滤波算法在厚尾非高斯噪声情况下的鲁棒性. 文献[25]同时考虑系统状态和噪声的协方差来构造基于混合统计相似性度量的代价函数,在此基础上提出了一种基于统计相似性度量框架的自适应离群鲁棒状态估计器,成功解决了异常离群点干扰下的状态估计问题. 文献[26]针对非高斯噪声下的传感器状态估计问题,利用多变量 Laplace 建模非高斯量测噪声,基于信息滤波的特点进一步提出了分布式多变量 Laplace 鲁棒状态估计算法. 文献[27]使用具有厚尾特性的学生氏 t 分布建模

输出噪声,提出了一种鲁棒高斯近似固定区间滤波器,适用于解决带有色厚尾量测噪声的非线性状态估计问题,并且比现有的方法具有更高的估计精度和更好的鲁棒性. 文献[28]通过建立基于学生氏 t 分布的鲁棒观测模型,在输出数据包含异常值的情况下提出了非线性状态空间系统的鲁棒建模方法,克服了异常值干扰下非线性过程建模带来的困难.

上述针对非高斯分布噪声设计的状态估计方法主要分为两个思路:(1)基于距离的方法. 此类算法均需要调节相应的超参数来保证算法估计性能. 以基于最大相关熵的滤波算法为例,该算法不针对特定类型的非高斯分布,但是其抑制非高斯噪声的能力与核宽度相关,同时缺乏严格的自适应更新依据,其估计精度提升有限^[29]. (2)基于分布的方法. 此类方法是通过数学上典型的厚尾分布来建模输出量测噪声,进而提升非线性系统状态估计算法的鲁棒性. 常用的典型厚尾分布有学生氏 t 分布、Laplace 分布、广义双曲分布等^[27,28,30-34]. 上述此类方法均假设状态转移噪声服从高斯分布,但在实际工业过程中由于传感器异常、现场设备故障等情况的出现不能保证状态转移过程服从高斯分布. 针对该问题文献[35]提出了更为一般性的方法,即非线性状态空间模型中的状态转移噪声与输出量测噪声可以遵循相互独立的任意分布. 该方法在考虑非高斯状态转移噪声的情况下进一步提升了模型的鲁棒性,然而依然存在以下两个问题:(1)典型的非高斯厚尾分布一般都有着复杂的数学表达式,直接将其应用于算法会增加计算复杂度,甚至会导致算法不存在解析解;(2)利用厚尾分布建模状态噪声和输出噪声虽然提高了算法鲁棒性,但是并不能从数学角度解释算法如何处理异常值.

针对状态噪声和输出噪声均为非高斯噪声的情况,本文提出了一种基于厚尾双学生氏 t 分布的非线性状态空间系统鲁棒辨识方法. 首先采用具有厚尾特性的学生氏 t 分布分别建模状态噪声和输出噪声,建立了非线性状态空间系统的概率形式模型. 其次,引入数学分解简化了学生氏 t 分布复杂的数学表达式,这样做的好处是:(1)降低了算法计算复杂度,抑制了异常值的负面影响;(2)从数学角度清晰地解释了本文所提算法中学生氏 t 分布对异常值的鲁棒性. 然后在期望最大化算法框架下实现未知参数估计. 最后利用数值例子和二连杆机械臂系统实例验证了本文所提出算法的有效性以及与现有方法相比的优越性.

2 问题描述及本文工作

2.1 问题描述

考虑如下非线性状态空间系统:

$$x_k = f(x_{k-1}, u_{k-1}, \theta_f) + w_{k-1} \quad (1)$$

$$y_k = g(x_k, \theta_g) + e_k \quad (2)$$

其中 $u_{1:N} = \{u_k\}_{k=1,2,\dots,N}$ 和 $y_{1:N} = \{y_k\}_{k=1,2,\dots,N}$ 分别表示输入和输出序列; $x_{1:N} = \{x_k\}_{k=1,2,\dots,N}$ 表示所有采样时刻的隐藏状态; $f(\cdot)$ 和 $g(\cdot)$ 分别表示非线性状态转移函数和非线性输出量测函数; θ_f 和 θ_g 表示为模型的参数; w_k 和 e_k 分别表示状态转移噪声和输出量测噪声, 它们相互独立并且可服从任意分布. 在多数现有的辨识方法中, 通常假设它们均遵循高斯分布^[16,35]. 然而在数据质量不理想的情况下, 例如存在异常离群点时, 基于高斯分布的噪声建模并不能保证辨识算法的鲁棒性. 图 1 展示了尺度参数和方差同为 1 时, 不同自由度条件下的学生氏 t 分布和高斯分布的对比结果.

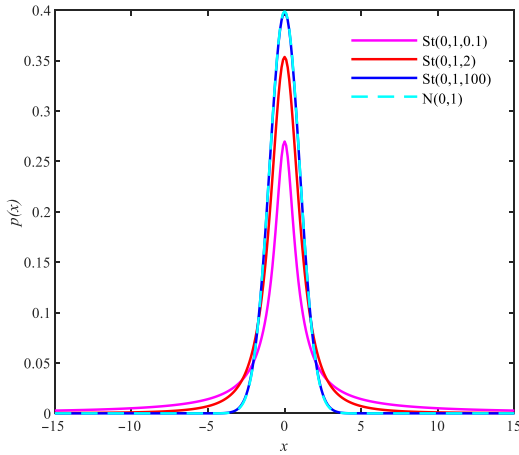


图 1 高斯分布和不同自由度参数下的学生 t 分布

从图 1 可以看出, 当自由度参数不断增大, 其尾部逐渐变短, 此时分布越发集中; 当自由度参数趋于无穷时, 学生氏 t 分布将会退化为高斯分布; 当学生氏 t 分布的自由度参数不断减小时, 其尾部越长, 这意味着此时学生氏 t 分布对离群值点的容忍度更高. 因此为了保证辨识算法的鲁棒性, 本文利用两个相互独立的学生氏 t 分布分别对状态转移噪声以及输出量测噪声建模如下:

$$w_k \sim \text{St}(w_k | 0, R_w, v_w) \quad (3)$$

$$e_k \sim \text{St}(e_k | 0, R_e, v_e) \quad (4)$$

其中, R_w 和 R_e 表示尺度参数, v_w 和 v_e 表示对应的自由度参数. 学生氏 t 分布的数学表达式如下^[28]:

$$\text{St}(w_k | 0, R_w, v_w) = \frac{\Gamma((v_w + 1)/2)(\pi R_w v_w)^{-1/2}}{\Gamma(v_w/2) \{1 + w_k^2 / (R_w v_w)\}^{(v_w + 1)/2}} \quad (5)$$

其中 $\Gamma(\cdot)$ 表示 Gamma 函数. 显然, 直接将式 (5) 应用于算法会导致其推导过程过于复杂. 因此基于学生氏 t 分布的数学性质, 通过引入辅助变量可将其做如下分解:

$$\text{St}(w_k | 0, R_w, v_w) = \int N(w_k | 0, R_w / \beta_k) p(\beta_k | v_w) d\beta_k \quad (6)$$

其中, β_k 表示未知的辅助变量; $N(w_k | 0, R_w / \beta_k)$ 表示方差随着 β_k 不断变化的子高斯分布; $p(\beta_k | v_w)$ 可看作每个子高斯分布的权重且服从以下伽玛分布^[28]:

$$p(\beta_k | v_w) = \text{Gamma}(\beta_k | v_w/2, v_w/2) \quad (7)$$

上述的数学分解方法将十分有助于本文算法的数学推导和实现. 基于式 (1) 和式 (5), 可得状态变量服从:

$$x_k \sim \text{St}(x_k | f(x_{k-1}, u_{k-1}, \theta_f), R_w, v_w) \quad (8)$$

通过数学分解有:

$$\begin{aligned} \text{St}(e_k | 0, R_e, v_e) &= \int N(e_k | 0, R_e / \lambda_k) p(\lambda_k | v_e) d\lambda_k \\ &= \frac{\Gamma((v_e + 1)/2)(\pi R_e v_e)^{-1/2}}{\Gamma(v_e/2) \{1 + e_k^2 / (R_e v_e)\}^{(v_e + 1)/2}} \quad (9) \end{aligned}$$

$$p(\lambda_k | v_e) = \text{Gamma}(\lambda_k | v_e/2, v_e/2)$$

同样的, 输出变量也服从学生氏 t 分布:

$$y_k \sim \text{St}(y_k | g(x_k, u_k, \theta_g), R_e, v_e) \quad (10)$$

2.2 本文工作

本文针对非线性状态空间系统中状态转移过程以及输出量测过程受到异常值干扰的问题, 利用两个相互独立的厚尾学生氏 t 分布对状态噪声和输出噪声进行建模, 提出了一种系统参数的鲁棒辨识算法. 在本文的工作中, 输入序列 $\{u_k\}_{k=1,2,\dots,N}$ 和输出序列 $\{y_k\}_{k=1,2,\dots,N}$ 用于构造可观数据集合 $C_{\text{obs}} = \{u_{1:N}, y_{1:N}\}$; 隐藏状态变量 $\{x_k\}_{k=1,2,\dots,N}$ 、辅助变量 $\{\beta_k\}_{k=1,2,\dots,N}$ 和辅助变量 $\{\lambda_k\}_{k=1,2,\dots,N}$ 用于构造隐藏变量数据集 $C_{\text{mis}} = \{x_{1:N}, \beta_{1:N}, \lambda_{1:N}\}$. 本文旨在数据集 C_{mis} 未知的情况下, 有针对性地解决数据异常问题, 通过观测数据集 C_{obs} 辨识未知参数集 $\Theta = \{\theta_f, \theta_g, R_w, v_w, R_e, v_e\}$.

本文的主要贡献总结如下:

(1) 利用学生氏 t 分布分别对状态噪声和输出噪声进行建模, 替换了传统辨识算法中的高斯分布, 保证了算法的鲁棒性;

(2) 通过将学生氏 t 分布分解为多个高斯分布的加权组合, 促进了算法的数学推导和实现, 为算法的鲁棒性提供了更清晰的解释;

(3) 同时得到了模型参数和噪声参数的迭代计算公式, 避免了分离独立计算带来的参数误差累积.

3 预备知识

3.1 期望最大化算法

在本文中, 由于隐藏变量数据集无法通过测量得到, 导致利用传统的极大似然估计来优化参数变得十分困难. 因此, 本文采用期望最大化 (Expectation Maximization, EM) 算法解决这个问题^[28]. EM 算法可以看作

是极大似然估计法的一个改进,在辨识数据集不完整或者包含隐藏变量的情况下,依然能够提供有效的参数估计. EM 算法本质上是一个迭代优化算法,通过重复执行期望步骤(E-step)和最大化步骤(M-step)直到获得未知的参数估计.

在 E-step 中,构造系统的对数似然函数 $\log p(C_{\text{mis}}, C_{\text{obs}}|\Theta)$, 求取其关于 C_{mis} 的期望, 得到代价函数(Q 函数):

$$Q(\Theta|\Theta^s) = E_{C_{\text{mis}}|C_{\text{obs}}, \Theta^s} \{ \log p(C_{\text{mis}}, C_{\text{obs}}|\Theta) \} \quad (11)$$

其中, Θ^s 表示第 s 次迭代估计的未知参数.

在 M-step 中,通过最大化 Q 函数,得到未知参数的第 $s+1$ 次估计为:

$$\Theta^{s+1} = \arg \max_{\Theta} Q(\Theta|\Theta^s) \quad (12)$$

通过不断迭代 E-step 和 M-step, 最终得到未知参数估计结果. 综上所述, EM 算法的核心步骤总结在算法 1 中.

算法 1 EM 算法

输入:可观测数据集 C_{obs} , 初始参数 Θ^0 , 最大迭代次数 S

1. E-step:根据式(11)构造代价函数
2. M-step:根据式(12)更新参数估计
3. 迭代终止:如果迭代次数 s 大于最大迭代次数 S 或者满足收敛条件,则迭代终止,否则返回步骤 1.

输出:参数集 Θ^*

3.2 粒子滤波

粒子滤波是一种计算非线性状态变量后验概率密度函数的有效方法,其核心思想为:通过一组有限个带有权重的粒子来模拟系统状态量的后验概率密度函数. 因此,在本文中状态变量的后验概率密度函数 $p(x_k|y_{1:k}, \Theta^s)$ 可以被计算为^[15]:

$$p(x_k|y_{1:k}, \Theta^s) \approx \sum_{l=1}^L W_k^l \delta(x_k - x_k^l) \quad (13)$$

其中, $l=1, 2, \dots, L$ 表示粒子数量; x_k^l 表示状态变量 x_k 的第 l 个粒子; W_k^l 表示第 l 个粒子对应的权重; $\delta(\cdot)$ 是狄拉克函数,它的性质如下:

$$\delta(x_k - x_k^l) = \begin{cases} 1, & x_k = x_k^l \\ 0, & \text{others} \end{cases} \quad (14)$$

同时,对于任意一个函数 $f(x)$, 下式积分可以被计算为:

$$\int_{-\infty}^{+\infty} f(x_k) \delta(x_k - x_k^l) dx_k = \sum_{l=1}^L f(x_k^l) \quad (15)$$

通常很难直接根据状态变量的后验概率密度函数直接采样获得粒子,这个问题常用的解决办法是引入重要性采样. 在引入重要性函数的基础上,通过对重要性函数进行采样以获得新的粒子. 此时,粒子的权重更新公式如下:

$$W_k^l \propto W_{k-1}^l \frac{p(y_k|x_k^l)p(x_k^l|x_{k-1}^l)}{q(x_k^l|y_{1:k}, x_{k-1}^l)} \quad (16)$$

其中 $q(\cdot)$ 表示重要性函数. 一般来说重要性密度函数通常选择具有先验性质的状态转移函数,因此有下式:

$$q(x_k|y_k, x_{k-1}) = p(x_k|x_{k-1}) \quad (17)$$

此时,每个粒子相对应的权重被更新为:

$$W_k^l \propto W_{k-1}^l p(y_k|x_k^l) \quad (18)$$

除此之外,还需要利用重采样技术避免出现粒子滤波中常见的粒子退化问题. 粒子退化问题指粒子权重的方差随着迭代次数增加而增大,即只有少数粒子权重很大,其他粒子权重近乎为零. 一般定义粒子多样性常数 N_{eff} 来确认退化现象是否严重,计算公式如下:

$$N_{\text{eff}} = \frac{1}{\sum_{l=1}^L (W_k^l)^2} \quad (19)$$

在每次粒子更新后,计算粒子多样性常数并且与算法设定的阈值 N_{Thd} 进行比较. 若 $N_{\text{eff}} < N_{\text{Thd}}$, 则说明粒子退化现象严重需要进行重采样. 经过重采样后,所有粒子的权重均等于 $1/L$.

综上,粒子滤波算法的核心步骤总结在算法 2 中.

算法 2 粒子滤波算法

输入:选取 L 个初始粒子 $\{x_0^l\}_{l=1,2,\dots,L}$ 并且设置对应权重为 $W_0^l = 1/L$,

输入状态变量的最大时刻 N , 多样性常数阈值 N_{Thd} , 设 $k=1$;

1. 根据重要性函数式(17)来更新粒子 $\{x_k^l\}_{l=1,2,\dots,L}$;
2. 根据式(18)更新粒子对应权重 $\{W_k^l\}_{l=1,2,\dots,L}$;
3. 根据式(19)计算粒子多样性常数,若 $N_{\text{eff}} < N_{\text{Thd}}$ 进行重采样步骤,否则进行下一步;
4. 令 $k=k+1$,重复执行步骤 1 至步骤 3,直至 $k=N$.

输出:1 到 N 时刻的粒子 $\{x_{1:N}^l\}_{l=1,2,\dots,L}$ 及其权重 $\{W_{1:N}^l\}_{l=1,2,\dots,L}$

3.3 粒子平滑

与粒子滤波算法相似,粒子平滑算法也提供了计算状态变量后验概率密度函数的有效方法^[16]. 首先平滑的概率密度函数 $p(x_k|y_{1:N}, \Theta^s)$ 进一步分解为:

$$p(x_k|y_{1:N}, \Theta^s) = p(x_k|y_{1:k}, \Theta^s) \times \int \frac{p(x_{k+1}|y_{1:N}, \Theta^s)p(x_{k+1}|x_k, \Theta^s)}{\int p(x_k|y_{1:k}, \Theta^s)p(x_{k+1}|x_k, \Theta^s) dx_k} dx_{k+1} \quad (20)$$

由上式可知,滤波函数 $p(x_k|y_{1:k}, \Theta^s)$ 、平滑概率分布 $p(x_{k+1}|y_{1:N}, \Theta^s)$ 以及状态转移分布 $p(x_{k+1}|x_k, \Theta^s)$ 被用于计算平滑的概率分布 $p(x_k|y_{1:N}, \Theta^s)$. 基于文献[16]中的结论,在时刻 k 处的平滑概率函数 $p(x_k|y_{1:N}, \Theta^s)$ 可以在数值上近似为:

$$p(x_k|y_{1:N}, \Theta^s) \approx \sum_{l=1}^L W_{k|N}^l \delta(x_k - x_k^l) \quad (21)$$

其中,相应的权重 $W_{k|N}^l$ 可以被计算为:

$$W_{k|N}^l = W_k^l \left[\frac{\sum_{j=1}^L W_{k+1|N}^j \frac{p(x_{k+1}^j | x_k^j, \Theta^s)}{\sum_{i=1}^L W_k^i p(x_{k+1}^i | x_k^i, \Theta^s)}}{\sum_{i=1}^L W_k^i p(x_{k+1}^i | x_k^i, \Theta^s)} \right] \quad (22)$$

综上,粒子平滑算法的核心步骤总结在算法3中.

算法3 粒子平滑算法

输入:算法2得到的粒子及其权重,初始化 $k=N$ 时刻的权重,并设 $k=N$;

1. 根据式(22)更新粒子对应的平滑权重 $\{W_{k|N}^l\}_{l=1,2,\dots,L}$;
2. 令 $k=k-1$,重复执行步骤1,直至 $k=1$.

输出:粒子平滑后的权重

4 基于厚尾双学生氏t分布的非线性状态空间系统鲁棒辨识算法

接下来将基于EM算法对本文所提出的鲁棒辨识算法进行详细推导.

4.1 E-step

根据算法1中的步骤2构造代价函数,首先计算所有数据的对数似然函数,基于概率链式法则可得:

$$\begin{aligned} \log p(C_{\text{mis}}, C_{\text{obs}} | \Theta) &= \log p(y_{1:N} | u_{1:N}, x_{1:N}, \beta_{1:N}, \lambda_{1:N}, \Theta) \\ &+ \log p(x_{1:N} | u_{1:N}, \lambda_{1:N}, \beta_{1:N}, \Theta) \\ &+ \log p(\lambda_{1:N} | u_{1:N}, \beta_{1:N}, \Theta) \\ &+ \log p(\beta_{1:N} | u_{1:N}, \Theta) + \log p(u_{1:N} | \Theta) \end{aligned} \quad (23)$$

结合式(1)~(10),对数似然函数可以进一步写为:

$$\begin{aligned} \log p(C_{\text{mis}}, C_{\text{obs}} | \Theta) &= \sum_{k=1}^N \log p(y_k | x_k, u_k, \lambda_k, \Theta) + \log p(x_1 | \Theta) \\ &+ \sum_{k=2}^N \log p(x_k | x_{k-1}, \beta_k, u_{k-1}, \Theta) \\ &+ \sum_{k=2}^N \log p(\beta_k | \Theta) + \sum_{k=1}^N \log p(\lambda_k | \Theta) \\ &+ \log p(u_{1:N} | \Theta) + \log p(\beta_1 | \Theta) \end{aligned} \quad (24)$$

对上式求取关于 $\log p(C_{\text{mis}}, C_{\text{obs}} | \Theta)$ 的条件期望,可以得到 Q 函数:

$$Q(\Theta | \Theta^s) = I_1 + I_2 + I_3 + C_1 \quad (25)$$

其中, I_1, I_2, I_3, C_1 分别表示如下:

$$\begin{aligned} I_1 &= \int p(x_1 | C_{\text{obs}}, \Theta^s) \times \left[\int p(\beta_1 | C_{\text{obs}}, \Theta^s) \times (\log p(x_1 | \Theta) \right. \\ &\quad \left. + \log p(\beta_1 | \Theta)) d\beta_1 \right] dx_1 \end{aligned} \quad (26)$$

$$\begin{aligned} I_2 &= \sum_{k=2}^N \iint p(x_k, x_{k-1} | C_{\text{obs}}, \Theta^s) \times \left[\int p(\beta_k | C_{\text{obs}}, x_k, x_{k-1}, \Theta^s) \right. \\ &\quad \left. \times (\log p(x_k | x_{k-1}, u_{k-1}, \beta_k, \Theta) + \log p(\beta_k | \Theta)) d\beta_k \right] dx_k dx_{k-1} \end{aligned} \quad (27)$$

$$\begin{aligned} I_3 &= \sum_{k=1}^N \int p(x_k | C_{\text{obs}}, \Theta^s) \times \left[\int p(\lambda_k | C_{\text{obs}}, \Theta^s) \times (\log p(\lambda_k | \Theta) \right. \\ &\quad \left. + \log p(y_k | x_k, u_k, \lambda_k, \Theta)) d\lambda_k \right] dx_k \end{aligned} \quad (28)$$

$$C_1 = \log p(u_{1:N} | \Theta) \quad (29)$$

根据上式易知, Q 函数中只有 I_2 和 I_3 用于对未知参数的估计.因此需要计算如下两项概率密度函数:

$$\begin{cases} p(x_k | C_{\text{obs}}, \Theta^s) \\ p(x_k, x_{k-1} | C_{\text{obs}}, \Theta^s) \end{cases} \quad (30)$$

式(30)是在已知全部观测数据的情况下求取状态变量的后验概率分布,可以看出这是一个平滑问题.由于本文考虑的是一般性的、非线性非高斯的状态平滑问题,因此采用算法3的粒子平滑算法来计算 $p(x_k | y_{1:N}, \Theta^s)$.相应地,联合后验概率分布 $p(x_k, x_{k+1} | C_{\text{obs}}, \Theta^s)$ 根据等式(20)分解为:

$$\begin{aligned} p(x_k, x_{k+1} | y_{1:N}, \Theta^s) &= p(x_k | y_{1:k}, \Theta^s) \\ &\quad \times \frac{p(x_{k+1} | y_{1:N}, \Theta^s) p(x_{k+1} | x_k, \Theta^s)}{\int p(x_k | y_{1:k}, \Theta^s) p(x_{k+1} | x_k, \Theta^s) dx_k} \end{aligned} \quad (31)$$

在时刻 k 处平滑的联合后验概率分布 $p(x_k, x_{k+1} | y_{1:N}, \Theta^s)$ 可以在数值上近似为:

$$p(x_k, x_{k+1} | y_{1:N}, \Theta^s) \approx \sum_{l=1}^L W_{k,k+1|N}^l \delta(x_k - x_k^l) \delta(x_{k+1} - x_{k+1}^l) \quad (32)$$

其中,相应的权重 $W_{k,k+1|N}^l$ 可以被计算为:

$$W_{k,k+1|N}^l = \frac{\kappa_k^l}{\sum_{i=1}^L \kappa_k^i} \quad (33)$$

其中,

$$\kappa_k^l = W_k^l W_{k+1|N}^l \frac{p(x_{k+1}^l | x_k^l, \Theta^s)}{\sum_{i=1}^L W_k^i p(x_{k+1}^i | x_k^i, \Theta^s)} \quad (34)$$

在通过粒子平滑算法求得上述两项概率密度函数后,根据式(15)、(21)、(22)以及(32)至(34), I_2 和 I_3 可以被计算为:

$$\begin{aligned} I_2 &\approx \sum_{k=2}^N \sum_{i=1}^L \sum_{j=1}^L W_{k,k-1|N}^{ij} \int p(\beta_k | C_{\text{obs}}, x_k, x_{k-1}, \Theta^s) \\ &\quad \times (\log p(x_k | x_{k-1}, u_{k-1}, \beta_k, \Theta) + \log p(\beta_k | \Theta)) d\beta_k \end{aligned} \quad (35)$$

$$\begin{aligned} I_3 &\approx \sum_{k=1}^N \sum_{i=1}^L W_{k|N}^i \int p(\lambda_k | C_{\text{obs}}, \Theta^s) \times (\log p(\lambda_k | \Theta) \\ &\quad + \log p(y_k | x_k, u_k, \lambda_k, \Theta)) d\lambda_k \end{aligned} \quad (36)$$

根据文[28]中的结论, I_2 可以被计算为:

$$\begin{aligned} I_2 &\approx \sum_{k=2}^N \sum_{i=1}^L \sum_{j=1}^L W_{k,k-1|N}^{ij} \left(\frac{v_w}{2} (\langle \log \beta_k \rangle^{ij} - \langle \beta_k \rangle^{ij}) \right. \\ &\quad \left. - \frac{1}{2} \langle \log \beta_k \rangle^{ij} - \frac{\langle \beta_k \rangle^{ij} (x_k^j - f(x_{k-1}^i, u_{k-1}, \theta_j))^2}{2R_w} \right) \\ &\quad + (N-1) \left(\frac{v_w}{2} \log \frac{v_w}{2} - \log \Gamma \left(\frac{v_w}{2} \right) - \frac{1}{2} \log 2\pi R_w \right) \end{aligned} \quad (37)$$

其中, $\Gamma(\cdot)$ 表示 Gamma 函数, $\langle \beta_k \rangle^{ij}$ 和 $\langle \log \beta_k \rangle^{ij}$ 分别表示 β_k 和 $\log \beta_k$ 的期望项, 通过下式计算:

$$\begin{aligned} \langle \beta_k \rangle^{ij} &= \int p(\beta_k | C_{\text{obs}}, x_k, x_{k-1}, \Theta^s) \beta_k d\beta_k \\ &= \frac{1 + v_w^s}{(x_k^j - f(x_{k-1}^i, u_{k-1}, \theta_f))^2 / R_w^s + v_w^s} \end{aligned} \quad (38)$$

$$\begin{aligned} \langle \log \beta_k \rangle^{ij} &= \int p(\beta_k | C_{\text{obs}}, x_k, x_{k-1}, \Theta^s) \log \beta_k d\beta_k \\ &= -\log\left(\frac{(x_k^j - f(x_{k-1}^i, u_{k-1}, \theta_f))^2 / R_w^s + v_w^s}{2}\right) \\ &\quad + \psi((1 + v_w^s)/2) \end{aligned} \quad (39)$$

其中, $\psi(x) = d \log(\Gamma(x)) / dx$ 表示 Digamma 函数. 同样, I_3 被计算为:

$$\begin{aligned} I_3 &\approx \sum_{k=1}^N \sum_{i=1}^L W_{k|N}^i \left(\frac{v_e}{2} (\langle \log \lambda_k \rangle^i - \langle \lambda_k \rangle^i) \right. \\ &\quad \left. - \frac{1}{2} \langle \log \lambda_k \rangle^i - \langle \lambda_k \rangle^i (y_k - g(x_k^i, u_k, \theta_g))^2 \right) \\ &\quad + N \left(\frac{v_e}{2} \log \frac{v_e}{2} - \frac{1}{2} \log 2\pi R_e - \log \Gamma\left(\frac{v_e}{2}\right) \right) \end{aligned} \quad (40)$$

其中, 期望项 $\langle \lambda_k \rangle^i$ 和 $\langle \log \lambda_k \rangle^i$ 通过下式计算可得:

$$\begin{aligned} \langle \lambda_k \rangle^i &= \int p(\lambda_k | C_{\text{obs}}, \Theta^s) \lambda_k d\lambda_k \\ &= \frac{1 + v_e^s}{(y_k - g(x_k^i, u_k, \theta_g))^2 / R_e^s + v_e^s} \end{aligned} \quad (41)$$

$$\begin{aligned} \langle \log \lambda_k \rangle^i &= \int p(\lambda_k | C_{\text{obs}}, \Theta^s) \log \lambda_k d\lambda_k \\ &= -\log\left(\frac{(y_k - g(x_k^i, u_k, \theta_g))^2 / R_e^s + v_e^s}{2}\right) \\ &\quad + \psi((1 + v_e^s)/2) \end{aligned} \quad (42)$$

4.2 M-step

根据 E-step 中构造的 Q 函数, 在 M 步中通过最大化 Q 函数对未知参数进行估计. 由上节可知, 在 Q 函数中只有 I_2 和 I_3 与参数有关. 因此令 I_2 和 I_3 关于相应变量的偏导数等于 0, 推导可得尺度参数 R_w 和 R_e 的估计为:

$$R_w^{s+1} = \frac{\sum_{k=2}^N \sum_{i=1}^L \sum_{j=1}^L W_{k,k-1|N}^{ij} \langle \beta_k \rangle^{ij} (x_k^j - f(x_{k-1}^i, u_{k-1}, \theta_f^s))^2}{N-1} \quad (43)$$

$$R_e^{s+1} = \frac{\sum_{k=1}^N \sum_{i=1}^L W_{k|N}^i \langle \lambda_k \rangle^i (y_k - g(x_k^i, u_k, \theta_g^s))^2}{N} \quad (44)$$

I_2 和 I_3 对自由度参数 v_w 求导后可得到下式:

$$\begin{aligned} &\frac{1}{2} \sum_{k=2}^N \sum_{i=1}^L \sum_{j=1}^L W_{k,k-1|N}^{ij} (\langle \log \beta_k \rangle^{ij} - \langle \beta_k \rangle^{ij}) \\ &\quad + \frac{N-1}{2} \left(\log \frac{v_w}{2} - \psi\left(\frac{v_w}{2}\right) + 1 \right) = 0 \end{aligned} \quad (45)$$

通过求解上述非线性方程可得到新的参数估计 v_w^{s+1} . 然而上式的闭式解通常不存在, 目前的文献中通常利用 Matlab 中的 fsolve 函数解决这个问题. 同样, 自

由度参数 v_e 可以通过求解下式得到:

$$\begin{aligned} &\frac{1}{2} \sum_{k=1}^N \sum_{i=1}^L W_{k|N}^i (\langle \log \lambda_k \rangle^i - \langle \lambda_k \rangle^i) + \frac{N}{2} \left(\log \frac{v_e}{2} - \psi\left(\frac{v_e}{2}\right) + 1 \right) \\ &= 0 \end{aligned} \quad (46)$$

由于在 I_2 中 θ_f 和 R_w 是相关的, 因此采用两步优化思想, 当更新 θ_f 时认为参数 R_w 是固定的. 所以, I_2 可以进一步简化为下式:

$$\text{sub}(I_2) = \sum_{k=2}^N \sum_{i=1}^L \sum_{j=1}^L W_{k,k-1|N}^{ij} \langle \beta_k \rangle^{ij} \times (x_k^j - f(x_{k-1}^i, u_{k-1}, \theta_f^s))^2 \quad (47)$$

参数 θ_f 可以通过最大化 $\text{sub}(I_2)$ 解决, 即:

$$\theta_f^{s+1} = \arg \max_{\theta_f} \text{sub}(I_2) \quad (48)$$

相似地, I_3 可以被简化为:

$$\text{sub}(I_3) = \sum_{k=1}^N \sum_{i=1}^L W_{k|N}^i \langle \lambda_k \rangle^i (y_k - g(x_k^i, u_k, \theta_g^s))^2 \quad (49)$$

因此, 参数 θ_g 同样可以通过最大化 $\text{sub}(I_3)$ 解决, 即:

$$\theta_g^{s+1} = \arg \max_{\theta_g} \text{sub}(I_3) \quad (50)$$

这里需要说明不能给出参数 θ_f 和 θ_g 解析解的原因, 这是因为它取决于特定的函数 $f(\cdot)$ 和 $g(\cdot)$.

综上所述, 本文提出的鲁棒辨识算法核心步骤总结在算法 4 中.

算法 4 双厚尾分布驱动的鲁棒辨识方法

输入: 辨识数据集 $C_{\text{obs}} = \{u_{1:N}, y_{1:N}\}$ 并初始化参数 $\Theta^s = \Theta^0$, 最大迭代次数 S

1. 根据式(21)和(32)分别计算状态后验分布 $p(x_k | y_{1:N}, \Theta^s)$ 以及联合状态后验分布 $p(x_k, x_{k+1} | y_{1:N}, \Theta^s)$;
2. 根据式(38)、(39)、(41)、(42)计算必要期望项 $\langle \beta_k \rangle^{ij}$ 、 $\langle \log \beta_k \rangle^{ij}$ 、 $\langle \lambda_k \rangle^i$ 和 $\langle \log \lambda_k \rangle^i$;
3. 根据式(43)和(44)更新尺度参数 R_w 和 R_e ;
4. 根据式(48)和(50)更新模型参数 θ_f 和 θ_g ;
5. 求解非线性方程式(45)和(46)得到新的自由度参数 v_w 和 v_e ;
6. 迭代终止: 如果迭代次数 s 大于最大迭代次数 S 或者满足收敛条件, 则迭代终止, 否则返回步骤 1.

输出: 输出辨识得到的参数集 $\hat{\Theta} = \{\hat{\theta}_f, \hat{\theta}_g, \hat{R}_w, \hat{R}_e, \hat{v}_w, \hat{v}_e\}$.

5 数值示例验证

考虑下面的非线性状态空间模型, 数学表达式为:

$$x_k = ax_{k-1} + b \frac{x_{k-1}}{1 + x_{k-1}^2} + u_k + w_{k-1} \quad (51)$$

$$y_k = cx_k^2 + e_k \quad (52)$$

针对上述数值例子, 本文所提算法可以写出模型参数估计的解析式. 首先根据式(47), 重写 $\text{sub}(I_2)$ 为:

$$\text{sub}(I_2) = \sum_{k=2}^N \sum_{i=1}^L \sum_{j=1}^L W_{k,k-1|N}^{ij} \times \langle \beta_k \rangle^{ij} (x_k^i - a x_{k-1}^i - b \frac{x_{k-1}^i}{1 + (x_{k-1}^i)^2} - u_k)^2 \quad (53)$$

此时可以得到参数 a 和 b 的估计公式为:

$$a_{\text{new}} = \frac{\sum_{k=2}^N \sum_{i=1}^L \sum_{j=1}^L W_{k,k-1|N}^{ij} \eta_k^{ij} x_{k-1}^i \langle \beta_k \rangle^{ij}}{\sum_{k=2}^N \sum_{i=1}^L \sum_{j=1}^L W_{k,k-1|N}^{ij} (x_{k-1}^i)^2 \langle \beta_k \rangle^{ij}} \quad (54)$$

$$b_{\text{new}} = \frac{\sum_{k=2}^N \sum_{i=1}^L \sum_{j=1}^L W_{k,k-1|N}^{ij} \zeta_k^{ij} \left(\frac{x_{k-1}^i}{1 + (x_{k-1}^i)^2} \right) \langle \beta_k \rangle^{ij}}{\sum_{k=2}^N \sum_{i=1}^L \sum_{j=1}^L W_{k,k-1|N}^{ij} \left(\frac{x_{k-1}^i}{1 + (x_{k-1}^i)^2} \right)^2 \langle \beta_k \rangle^{ij}} \quad (55)$$

其中,

$$\eta_k^{ij} = x_k^i - b_{\text{old}} \frac{x_{k-1}^i}{1 + (x_{k-1}^i)^2} - u_k \quad (56)$$

$$\zeta_k^{ij} = x_k^i - a_{\text{old}} x_{k-1}^i - u_k \quad (57)$$

同样根据式(49),重写 $\text{sub}(I_3)$ 为:

$$\text{sub}(I_3) = \sum_{k=1}^N \sum_{i=1}^L W_{k|N}^i \langle \lambda_k \rangle^i (y_k - c(x_k^i)^2)^2 \quad (58)$$

因此,可以得到参数 c 的估计公式为:

$$c_{\text{new}} = \frac{\sum_{k=1}^N \sum_{i=1}^L W_{k|N}^i y_k (x_k^i)^2 \langle \lambda_k \rangle^i}{\sum_{k=1}^N \sum_{i=1}^L W_{k|N}^i (x_k^i)^4 \langle \lambda_k \rangle^i} \quad (59)$$

5.1 鲁棒性测试

在本节中,为了测试算法对异常值的鲁棒性,将均匀分布在 $[-5, 5]$ 范围内的数据当成异常值并与高斯噪声混合,分别加入到状态更新过程和输出量测过程中. 这样做的目的是检测算法能否对包含异常值的混合噪声进行精准建模并同时给出精确的模型参数估计. 测试环境具体设置如下:

(1) 模型真实参数设置为 $a=0.6$ 、 $b=2$ 以及 $c=0.5$, 算法迭代次数设置为 200 次.

(2) 输入信号设置为 $u_k = 4\cos(1.2k)$, 在此基础上收集 $N=500$ 个辨识数据. 图 2 中展示了异常值比例为 10% 时的混合噪声数据.

(3) 算法 4 在辨识实验中初始值设为 $a_0=0.3$ 、 $b_0=1.2$ 、 $c_0=0.2$.

(4) 仿真实验采用 MATLAB R2022a 版本软件进行,所用计算机的 CPU 型号为英特尔酷睿 i7-7700,主频为 3.60 GHz,内存为 16 GB.

在上述设置下,为了评估本文所提出算法的鲁棒性,考虑如下 2 种仿真:

实验 1 测试算法 4 在异常值比例为 10% 时的鲁

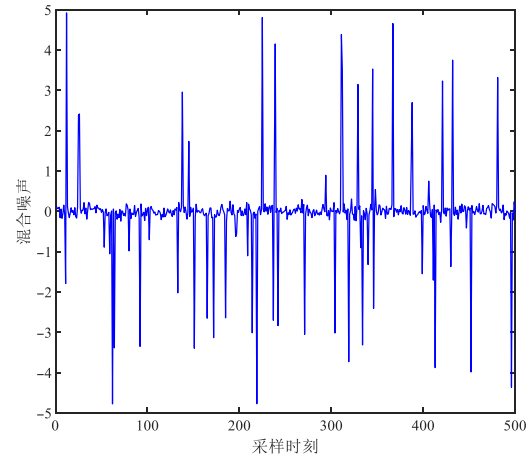


图 2 混合噪声

棒性.

图 3 展示了在状态过程和输出过程均受到 10% 异常值干扰时,系统参数的收敛曲线. 不难看出,在大约迭代 100 次后所有参数均能精确地收敛到真值. 这说明本文提出的算法 4 在状态转移过程以及输出量测过程受到异常值干扰的情况下,依然能精确地辨识出模型参数,有效地处理混合噪声中的异常值.

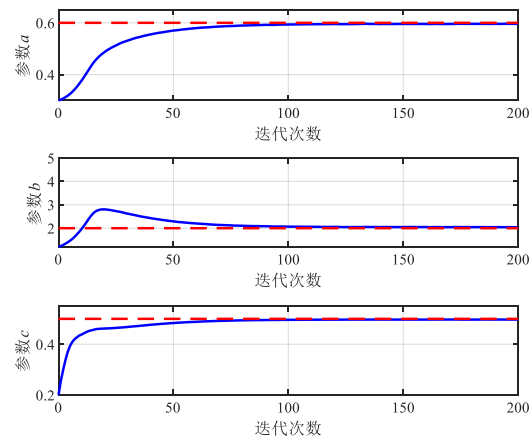


图 3 异常值比例为 10% 的模型参数收敛曲线

图 4 展示了无噪声的状态变量、输出变量和模型预测的状态变量、输出变量的对比结果. 不难看出,由于算法 4 得到了精确的模型参数估计,使得模型预测的状态变量、输出变量能够精准的拟合真实值,这说明了算法 4 有着较高的模型预测精度.

图 5 展示了 Q 函数随着模型参数 a 和 c 的三维增长曲线,其中红色曲线为参数 a 和 c 的估计轨迹. 显然 Q 函数不断增长并最终迭代至稳态值,此时得到了参数 a 和 c 的估计值,这也验证了算法 4 的收敛性.

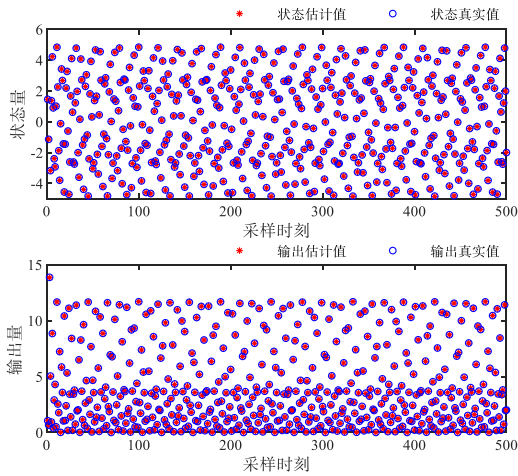


图4 异常值比例为10%的系统状态估计和输出估计

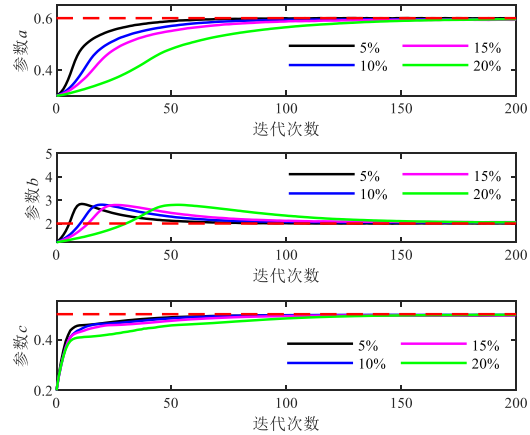


图6 异常值比例不同时参数估计结果

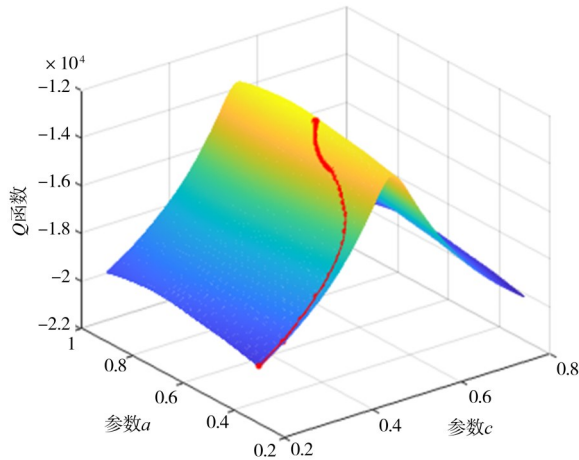


图5 Q 函数关于参数 a 和 c 的三维图

实验 2 比较不同异常值比例条件下算法 4 的辨识结果.

图 6 分别展示了在异常值比例为 5%、10%、15% 和 20% 时, 参数估计的比较结果. 从图中结果可以看出, 在不同的数据质量条件下, 本文所提算法 4 估计的参数最终都能收敛到真值附近, 并且随着异常值百分比减小, 参数的收敛速度逐渐加快.

通过上述两个仿真可以得出算法 4 能够有效的处理异常值, 同时有着较高的参数估计精度和模型预测精度. 且在异常值比例增加时, 算法 4 也依旧保持了较强的鲁棒性.

5.2 对比试验

为了进一步验证算法的有效性, 将本文提出的算法与同类算法进行比较. 对比算法选取如下:

(1) 在文献[16]中, Gopaluni 提出了一种基于高斯分布的非线性状态空间系统辨识方法, 记为 Gaussian-

Iden;

(2) 在文献[28]中, Liu 等人基于学生氏 t 分布提出了一种仅考虑了输出异常值的鲁棒辨识算法, 状态转移过程仍为高斯过程, 记为 Student-Single;

(3) 在文献[35]中, Schön 等人提出了一种非线性状态空间系统辨识方法, 其中状态噪声和输出噪声可以服从任意分布. 但是该方法直接应用噪声分布的数学表达式, 当状态噪声和输出噪声服从学生氏 t 分布时, 模型参数不存在解析解, 需采用阻尼牛顿法来迭代寻找模型参数, 记为 Student-Newton.

在对比实验中, 通过以下两个指标来量化各算法的参数估计精度与模型预测效果, 分别是均方根误差 (Root Mean Square Error, RMSE) 和相对误差参数 (Relative Parameter Estimation Error, RPEE), 定义如下:

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n^* - \hat{y}_n)^2} \quad (60)$$

$$RPEE = \frac{\|\hat{\theta} - \theta^{true}\|_2}{\|\theta^{true}\|_2} \times 100\% \quad (61)$$

其中, N 为样本数, y_n^* 和 \hat{y}_n 分别为第 n 个采样时刻的理想输出及估计输出, θ^{true} 和 $\hat{\theta}$ 分别是参数的真实值和估计值.

实验 3 状态过程和输出过程均受 10% 的异常值干扰时参数辨识精度对比.

图 7 展示了在 10% 的异常值条件下, 各算法的参数辨识对比结果. 图 8 给出了各算法在每次迭代时的 RMSE 结果对比. 从上述辨识结果不难看出:

(1) 从图 7 可以看出, 当状态噪声和输出噪声包含异常值时, 即系统噪声不再服从高斯分布时, Gaussian-Iden 和 Student-Single 方法的参数辨识精度将大幅度下降, Student-Single 方法的参数辨识精度略高于 Gaussian-

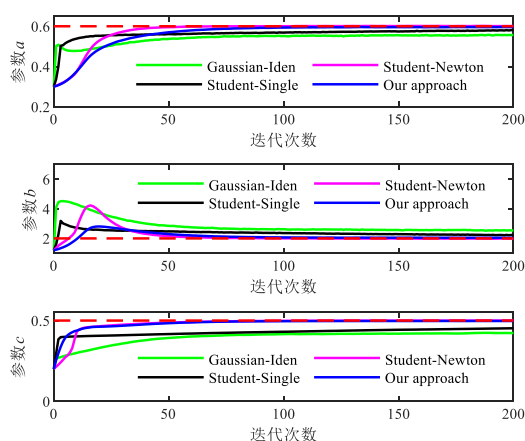


图7 异常值比例为10%的比较结果

Iden方法. 这说明当高斯分布不足以描述噪声的统计特性时,模型参数的辨识精度将无法得到保障.

(2)从图7和8可以看出,算法4与Student-Newton方法均得到了精度较高的参数辨识结果,这是因为这两种方法均假设状态噪声和输出噪声服从学生氏t分布.与高斯分布相比,学生氏t分布有更长的尾部对异常值的容忍度更高,通过调节学生氏t分布的超参数能够自适应改变其统计特性,并更加精准地描述噪声的统计特性,保障了算法对于异常值的鲁棒性.

关于Student-Newton方法与本文提出的算法4的区别将在下一节详细介绍.

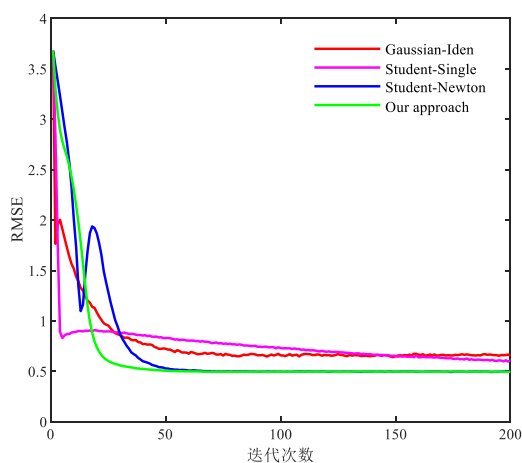


图8 异常值比例为10%时不同方法每次迭代的RMSE值

实验4 蒙特卡洛对比实验.

为了进一步验证本文提出的算法4的稳定性和有效性,本文设计了蒙特卡洛对比实验.在每次蒙特卡洛实验中,独立运行每个算法100次.在蒙特卡洛实验中,参数初始值 a_0 、 b_0 、 c_0 分别从 $[0.3\ 0.5]$ 、 $[1.0\ 1.4]$ 、

$[0.2\ 0.4]$ 中随机选择.所得到的蒙特卡洛实验结果如表1和表2所示,表1展示了当异常值比例为10%的情况下,各方法在蒙特卡洛实验中辨识得到的参数估计均值、标准差以及算法运行时间;表2展示了在不同异常值比例条件下,各方法在蒙特卡洛实验中平均RMSE和RPEE值.根据上述对比实验结果,可以得出以下结论:

(1)从表1可以看出,Gaussian-Iden方法和Student-Single方法的参数估计均值和真实值有较大的误差,但是算法的运行时间相对较少,证明算法的复杂度相对较低.

(2)从表1和表2可以看出,算法4的参数估计精度和Student-Newton方法比较近似,在不同数据条件下均能准确地估计出模型参数的真实值.但是由于Student-Newton方法中采用了解析法搜寻模型参数,使得算法的计算复杂度比本文提出的算法4要高.

(3)从表1和表2可以看出,在不同的异常值条件下以及在不同的初始值条件下,本文所提算法4均能准确地估计出模型参数,这验证了算法的稳定性和有效性.

(4)从表2可以看出,当异常值比例不断增加时,算法4得到的RPEE值不断增加.这说明算法4的表现与数据质量息息相关,当异常值比例不断增加时,数据质量不断降低,此时算法4的参数估计精度也在下降.

5.3 实验讨论

在上述算法验证过程中,本文所提算法4与Student-Newton方法的参数辨识精度相当.与Student-Newton方法相比,本文的算法4主要优势体现在以下几个地方:

(1) 算法4针对异常值具有鲁棒性的理论解释更加清楚.

Student-Newton方法采用阻尼牛顿法优化代价函数,搜索寻找最优参数,但这并不能清楚地解释为何算法能够处理数据中的异常值.而本文提出的算法4可以清楚地解释为何算法对异常值具有鲁棒性.以模型参数 c 为例,其估计表达式如式(59)所示.其中 $\langle \lambda_k \rangle^i$ 可以视为一个权重由算法自适应地分配给每一个输出变量 y_k ,结合式(41)和(52), $\langle \lambda_k \rangle^i$ 的具体表达式可改写如下:

$$\langle \lambda_k \rangle^i = \int p(\lambda_k | C_{\text{obs}}, \Theta^s) \lambda_k d\lambda_k = \frac{1 + v_e^s}{(y_k - c(x_k^i)^2)^2 / R_e^s + v_e^s} \quad (62)$$

由上式可以看出,权重 $\langle \lambda_k \rangle^i$ 与预测输出和实际输出的误差平方成反比.假设在某个时刻输出数据点 y_k 是异常值,那么在这个时刻输出误差 $(y_k - c(x_k^i)^2)$ 应该相对较大,这就意味着分配给输出数据 y_k 的权重相对较小,甚至趋近于0.此时,异常值对参数估计的影响被大大降低,这确保了算法对异常值具有鲁棒性.

(2) 算法 4 的计算效率更高.

以尺度参数 R_e 的求解为例, 在 Student-Newton 方法中需要通过优化下面的代价函数:

$$\begin{aligned}
 I_3^n &= \sum_{i=1}^N \sum_{k=1}^L W_{k|N}^i \log p(y_k | x_k^i, \theta) \\
 &= \sum_{i=1}^N \sum_{k=1}^L W_{k|N}^i \log \left\{ \frac{\Gamma((1+\nu_e)/2)(\pi\nu_e R_e)^{-1/2}}{\Gamma(\nu_e/2)} \right. \\
 &\quad \left. \times (1 + (y_k - c(x_k^i))^2 / (\nu_e R_e))^{-(1+\nu_e)/2} \right\} \quad (63)
 \end{aligned}$$

显然由于在 Student-Newton 方法中直接应用了学生氏 t 分布的数学表达式, 使得代价函数过于复杂, 使得 R_e 的解析解不存在. 因此需要使用额外的数学优化手段来搜寻 R_e 的估计值, 这无疑会增加算法的复杂度.

在本文的算法 4 中, 尺度参数 R_e 的解析解由式 (44) 给出. 此外, 表 2 中算法的计算时间也表明, 算法 4 的计算效率更高.

表 1 异常值比例为 10% 时蒙特卡洛仿真参数估计结果的均值和标准差

方法	a		b		c		时间/s
	均值	标准差	均值	标准差	均值	标准差	
Gaussian-Iden	0.555 8	0.001 1	2.551 1	0.014 8	0.421 9	0.001 5	486
Student-Single	0.583 2	0.002 8	2.188 2	0.032 5	0.458 8	0.005 7	1 227
Student-Newton	0.598 9	0.000 3	2.005 9	0.059 5	0.498 5	0.000 6	13 410
Our approach	0.597 5	0.002 0	2.000 3	0.009 3	0.492 9	0.000 2	10 015

表 2 异常值比例不同时蒙特卡洛仿真的平均 RMSE、RPEE 值

比例	Gaussian-Iden		Student-Single		Student-Newton		Our approach	
	RMSE	RPEE	RMSE	RPEE	RMSE	RPEE	RMSE	RPEE
5%	0.576 6	9.27%	0.524 0	4.68%	0.489 9	0.23%	0.484 2	0.32%
10%	0.775 3	26.00%	0.568 5	9.01%	0.495 1	0.29%	0.481 5	0.52%
15%	1.075 4	38.37%	0.748 0	11.24%	0.504 9	2.20%	0.494 8	2.17%
20%	1.839 9	79.51%	0.910 4	26.24%	0.511 9	2.82%	0.510 9	3.56%

6 二连杆机械臂系统验证

二连杆机械臂是现实工业中常见的一种刚性机械系统, 其机械简图如图 9 所示. 其中, $F(t)$ 表示垂直向下的外力, 即角度 $\alpha_3 = 90^\circ$. 其他物理变量的解释见文献 [13, 36]. 假设二连杆机械臂系统受到一个很小的随机外界干扰, 那么它的动态特性可以数学表示为^[13, 36]

$$\mathbf{M}\ddot{\Phi} + \mathbf{E}\dot{\Phi} + \mathbf{K}(t)\Phi = 0 \quad (64)$$

其中,

$$\mathbf{M} = \begin{bmatrix} b_1 & 0 \\ 0 & b_3 \end{bmatrix}, \mathbf{E} = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}, \mathbf{K}(t) = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \quad (65)$$

$$\begin{aligned}
 E_{11} &= d_1 + d_2, E_{12} = E_{21} = -d_2, E_{22} = d_2, \\
 K_{11} &= k_1 + k_2 - b_4 \sin(\alpha_{10}) - F(t)l \cos(\alpha_{10} - \alpha_3), \\
 K_{22} &= k_2 - b_5 \sin(\alpha_{20}) - F(t)l \cos(\alpha_{20} - \alpha_3), \quad (66)
 \end{aligned}$$

$$\begin{aligned}
 K_{12} &= K_{21} = -k_2, \Phi = [\alpha_{11} \quad \alpha_{21}]^T, \\
 b_1 &= 4ml^2/3, b_2 = ml^2/2, b_3 = ml^2/3, \\
 b_4 &= 3ml^2 g/2, b_5 = ml^2 g/2
 \end{aligned}$$

选取状态变量为 $\mathbf{x} = [\dot{\Phi}^T \quad \Phi^T]^T$, 则可将式 (64) 写为:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\mathbf{x} \quad (67)$$

其中,

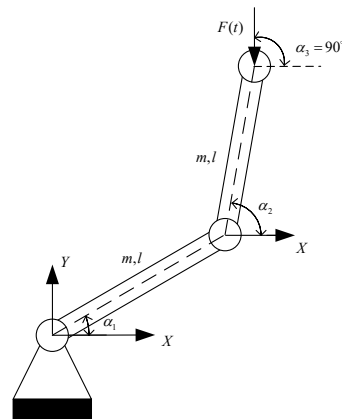


图 9 二连杆机械臂系统简图

$$\mathbf{A}(t) = \begin{bmatrix} -\frac{E_{11}}{b_1} & -\frac{E_{12}}{b_1} & -\frac{K_{11}}{b_1} & -\frac{K_{12}}{b_1} \\ -\frac{E_{21}}{b_3} & -\frac{E_{22}}{b_3} & -\frac{K_{21}}{b_3} & -\frac{K_{22}}{b_3} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (68)$$

利用有限差分离散化方法可将式 (67) 和 (68) 中的连续状态空间模型离散化为非线性状态空间模型, 其中状态转移函数和输出量测函数分别为^[13, 36]

$$f(\mathbf{x}_{k-1}, u_{k-1}, \theta_f) = \mathbf{x}_{k-1} + \Delta t \begin{bmatrix} -\frac{E_{11}}{b_1} & -\frac{E_{12}}{b_1} & -\frac{K_{11}}{b_1} & -\frac{K_{12}}{b_1} \\ -\frac{E_{21}}{b_3} & -\frac{E_{22}}{b_3} & -\frac{K_{21}}{b_3} & -\frac{K_{22}}{b_3} \\ \theta_1 & 0 & 0 & 0 \\ 0 & \theta_2 & 0 & 0 \end{bmatrix} \mathbf{x}_{k-1} \quad (69)$$

$$\mathbf{g}(\mathbf{x}_k, \theta_g) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k \quad (70)$$

其中, Δt 为采样周期, θ_1 和 θ_2 为待估计的模型参数, 真实值设为 $\theta_1 = \theta_2 = 1$. 其他参数的取值见文献[13].

在本节的实验中, 选取周期信号 $F(t) = 10\sin(0.02\pi t)$ 为输入信号并用其来激励系统, 采集 $N = 500$ 个输出数据. 为了验证算法4的鲁棒性能, 在数据中加入不同比例的异常值来降低辨识数据的质量, 当异常值比例为10%时的数据如图10所示.

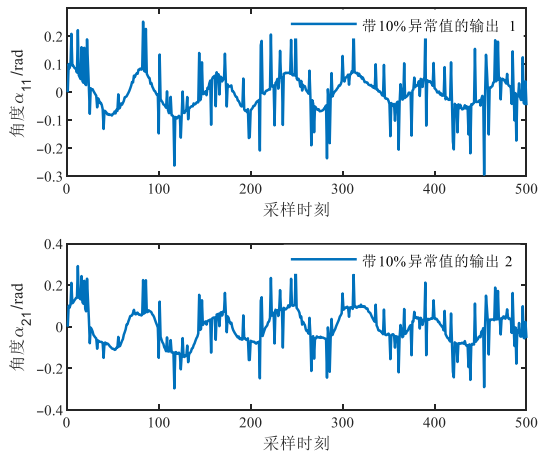


图10 异常值比例为10%的二连杆机械臂系统输出数据

实验5 不同异常值比例条件下算法4的验证结果.

为了验证本文所提算法4在具体实例上的辨识效果, 图11分别展示了在异常值比例为0%、5%、10%和15%时参数估计的比较结果. 从图11中的辨识结果可以看出, 经过200次迭代后, 当辨识数据集被不同百分比的异常值破坏时, 算法4都能够精确地估计二连杆机械臂系统模型参数.

实验6 异常值比例为10%的条件下对比实验验证结果.

图12展示了在10%的异常值条件下各算法的参数辨识对比结果. 从上述结果中可以看出:

(1) 基于算法4和 Student-Newton 方法得到的参数估计值能稳定地收敛到真实值, 且算法4的参数收敛速

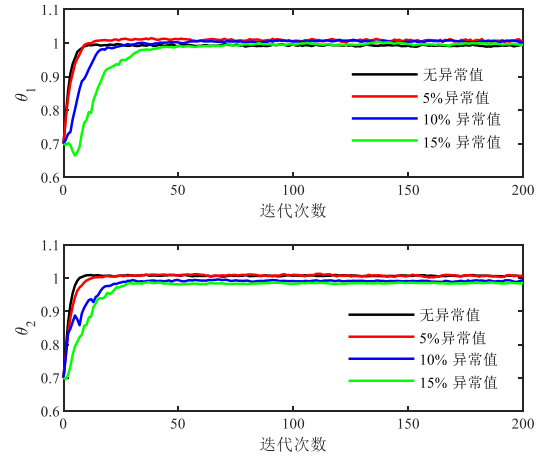


图11 异常值比例不同时参数 θ_1 和 θ_2 估计结果

度比 Student-Newton 方法快.

(2) 基于 Gaussian-Iden 和 Student-Single 算法得到的参数估计值均呈现出波动状态, 无法稳定的收敛到真实值, 甚至出现了有偏的参数估计.

基于以上两个二连杆机械臂系统验证实验, 可以看出算法4能够有效地解决异常值对模型参数估计的影响. 即在异常值干扰条件下, 算法4能够得到准确的模型参数估计, 这证明了本文所提算法4的有效性.

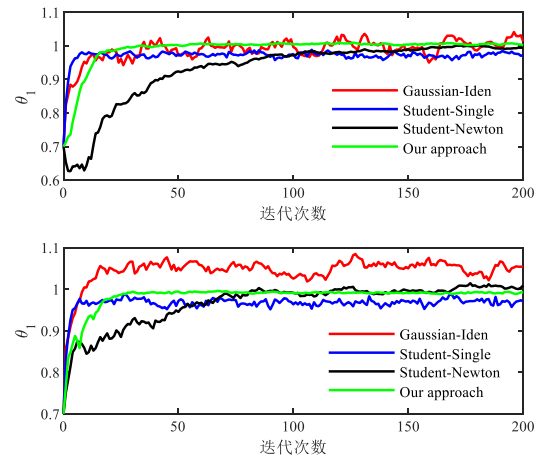


图12 异常值比例为10%时不同方法的参数 θ_1 和 θ_2 估计结果

7 结论

本文主要针对非线性状态空间系统辨识问题开展研究, 解决了状态转移过程和输出量测过程均受异常值干扰的问题, 提出了一种更加全面的鲁棒辨识算法. 通过利用两个相互独立的厚尾学生氏 t 分布分别对状态噪声和输出噪声进行建模, 保障了算法的鲁棒性; 利用学生氏 t 分布的数学分解为算法的鲁棒性提供了更加清晰的解释, 也进一步提高了算法的计算效率. 最后

通过对比实验验证算法的稳定性和鲁棒性。

参考文献

- [1] 刘鑫. 时滞取值概率未知下的线性时滞系统辨识方法[J]. 自动化学报, 2023, 49(10): 2136-2144.
LIU X. Identification of linear time-delay systems with unknown delay distributions in its value range[J]. Acta Automatica Sinica, 2023, 49(10): 2136-2144. (in Chinese)
- [2] DING F, LIU X M, HAYAT T. Hierarchical least squares identification for feedback nonlinear equation-error systems[J]. Journal of the Franklin Institute, 2020, 357(5): 2958-2977.
- [3] HAN R J, SALEHI Y, HUANG B, et al. Parameter estimation for nonlinear systems with multirate measurements and random delays[J]. AIChE Journal, 2021, 67(9): 17327-1-17327-15.
- [4] LIU X, YANG X Q, LIU X F. Nonlinear state-space system identification with robust Laplace model[J]. International Journal of Control, 2021, 94(6): 1492-1501.
- [5] LIU Q, LIN W Y, JIANG S L, et al. Robust estimation of Wiener models in the presence of outliers using the VB approach[J]. IEEE Transactions on Industrial Electronics, 2021, 68(11): 11390-11399.
- [6] HOU J, SU H, YU C P, et al. Bias-correction errors-in-variables Hammerstein model identification[J]. IEEE Transactions on Industrial Electronics, 2023, 70(7): 7268-7279.
- [7] PODUSENKO A, AKBAYRAK S, SENZO I, et al. Message passing-based system identification for NARMAX models[C]//2022 IEEE 61st Conference on Decision and Control (CDC). Piscataway: IEEE, 2022: 7309-7314.
- [8] GU Y, ZHU Q M, NOURI H. Identification and U-control of a state-space system with time-delay[J]. International Journal of Adaptive Control and Signal Processing, 2022, 36(1): 138-154.
- [9] ZOU Y D, HU W Q, XIAO Z H, et al. Design of intelligent nonlinear robust controller for hydro-turbine governing system based on state-dynamic-measurement hybrid feedback linearization method[J]. Renewable Energy, 2023, 204: 635-651.
- [10] 张宏伟. 时空学习驱动的混合重要性高斯滤波[EB/OL]. (2023-09-26)[2023-11-16]. <https://doi.org/10.12263/DZXB.20230056>.
- [11] 奚畅, 蔡志明, 袁骏. 被动声纳粒子滤波检测前跟踪的检测门限设置方法[J]. 电子学报, 2022, 50(7): 1722-1734.
XI C, CAI Z M, YUAN J. A method of determining detection threshold for particle filter track-before-detect in passive sonar[J]. Acta Electronica Sinica, 2022, 50(7): 1722-1734. (in Chinese)
- [12] BRUNOT M, JANOT A, CARRILLO F, et al. State space estimation method for robot identification[J]. Ifac Papersonline, 2016, 49(21): 228-233.
- [13] LIU K. Identification of linear time-varying systems[J]. Journal of Sound Vibration, 1997, 206(4): 487-505.
- [14] GHIAUS C, GHAZAL R, JOUBERT P, et al. Gray-box state-space model and parameter identification of desiccant wheels[J]. Applied Thermal Engineering, 2013, 51(1/2): 742-752.
- [15] LIU X P, YANG X Q. Identification of nonlinear state-space systems with skewed measurement noises[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2022, 69(11): 4654-4662.
- [16] GOPALUNI R B. A particle filter approach to identification of nonlinear processes under missing observations[J]. The Canadian Journal of Chemical Engineering, 2008, 86(6): 1081-1092.
- [17] LIU X T, LI Q L, WANG L, et al. Data-driven state of charge estimation for power battery with improved extended Kalman filter[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-10.
- [18] 陈云翔, 王泽洲, 蔡忠义, 等. 基于EM-EKF与隐含比例退化模型的机载电子设备剩余寿命自适应预测[J]. 电子学报, 2021, 49(3): 500-509.
CHEN Y X, WANG Z Z, CAI Z Y, et al. Adaptive prediction of remaining useful lifetime for the airborne electronic equipment based on the EM-EKF algorithm and hidden degradation model with the proportion relationship[J]. Acta Electronica Sinica, 2021, 49(3): 500-509. (in Chinese)
- [19] ZHU F, FU J Q. A novel state-of-health estimation for lithium-ion battery via unscented Kalman filter and improved unscented particle filter[J]. IEEE Sensors Journal, 2021, 21(22): 25449-25456.
- [20] LI G H, NIU D B, SONG E B. A multiple-model state estimator based on the unscented Kalman filter[C]//2020 39th Chinese Control Conference (CCC). Piscataway: IEEE, 2020: 3011-3016.
- [21] 卢春光, 张永顺, 李志汇, 等. 量测随机延迟下带厚尾噪声的鲁棒 Student's t 随机容积分卡尔曼滤波器[J]. 电子学报, 2019, 47(9): 1830-1840.
LU C G, ZHANG Y S, LI Z H, et al. Robust Student's t

- based stochastic cubature Kalman filter with heavy-tailed noises and randomly delayed measurements[J]. *Acta Electronica Sinica*, 2019, 47(9): 1830-1840. (in Chinese)
- [22] SWATI. Continuous discrete cubature quadrature Kalman filter[J]. *Asian Journal of Control*, 2022, 24(1): 483-493.
- [23] GUO F, HARIPRASAD K, HUANG B, et al. Robust identification for nonlinear errors-in-variables systems using the EM algorithm[J]. *Journal of Process Control*, 2017, 54: 129-137.
- [24] CHEN B D, DANG L J, GU Y T, et al. Minimum error entropy Kalman filter[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, 51(9): 5819-5829.
- [25] BAI M M, HUANG Y L, ZHANG Y G, et al. Statistical similarity measure-based adaptive outlier-robust state estimator with applications[J]. *IEEE Transactions on Automatic Control*, 2022, 67(8): 4354-4361.
- [26] 王国庆, 杨春雨, 马磊. 基于多变量 Laplace 分布的非线性系统分布式鲁棒状态估计[J]. *电子学报*, 2022, 50(11): 2610-2618.
WANG G Q, YANG C Y, MA L. Distributed robust state estimation for nonlinear systems based on multivariate Laplace distribution[J]. *Acta Electronica Sinica*, 2022, 50(11): 2610-2618. (in Chinese)
- [27] 黄玉龙, 张勇刚, 李宁, 等. 一种带有色量测噪声的非线性系统辨识方法[J]. *自动化学报*, 2015, 41(11): 1877-1892.
HUANG Y L, ZHANG Y G, LI N, et al. An identification method for nonlinear systems with colored measurement noise[J]. *Acta Automatica Sinica*, 2015, 41(11): 1877-1892. (in Chinese)
- [28] LIU X, YANG X Q. Robust identification approach for nonlinear state-space models[J]. *Neurocomputing*, 2019, 333: 329-338.
- [29] CHEN B D, LIU X, ZHAO H Q, et al. Maximum correntropy Kalman filter[J]. *Automatica*, 2017, 76: 70-77.
- [30] SAMMAKNEJAD N, ZHAO Y J, HUANG B. A review of the expectation maximization algorithm in data-driven process identification[J]. *Journal of Process Control*, 2019, 73: 123-136.
- [31] WANG D P, ZHANG H, HUANG H L. An adaptive robust student's t-based Kalman filter based on multi-sensor fusion[C]//15th International Conference on Intelligent Robotics and Applications (ICIRA). Cham: Springer International Publishing, 2022: 603-613.
- [32] YANG X Q, YIN S. Robust global identification and output estimation for LPV dual-rate systems subjected to random output time-delays[J]. *IEEE Transactions on Industrial Informatics*, 2017, 13(6): 2876-2885.
- [33] YANG X Q, LIU X, HAN B X. LPV model identification with an unknown scheduling variable in the presence of missing observations-a robust global approach[J]. *IET Control Theory & Applications*, 2018, 12(10): 1465-1473.
- [34] 王国庆, 杨春雨, 马磊, 等. 基于高斯-广义双曲混合分布的非线性卡尔曼滤波[J]. *自动化学报*, 2023, 49(2): 448-460.
WANG G Q, YANG C Y, MA L, et al. Nonlinear Kalman filter based on gaussian-generalized-hyperbolic mixing distribution[J]. *Acta Automatica Sinica*, 2023, 49(2): 448-460. (in Chinese)
- [35] SCHÖN T B, WILLS A, NINNESS B. System identification of nonlinear state-space models[J]. *Automatica*, 2011, 47(1): 39-49.
- [36] JABALI M B A, KAZEMI M H. Uncertain polytopic LPV modelling of robot manipulators and trajectory tracking[J]. *International Journal of Control, Automation and Systems*, 2017, 15(2): 883-891.

作者简介



刘鑫 男, 1990年7月出生于江苏省盐城市. 现为中国矿业大学人工智能研究院副教授、硕士生导师. 从事系统辨识、数据驱动的工业建模和软测量方面的研究.
E-mail: 15B904027@hit.edu.cn



海洋 男, 1999年5月出生于辽宁省阜新市. 现为中国矿业大学信息与控制工程学院硕士研究生. 主要研究方向为系统辨识, 数据驱动的工程建模.
E-mail: ts22060067a31@cumt.edu.cn



代伟 男, 1984年3月出生于河南省安阳市. 现为现为中国矿业大学信息与控制工程学院教授、博士生导师. 从事复杂工业过程建模、运行优化与控制方面的研究.
E-mail: weidai@cumt.edu.cn