

基于近似偏最小一乘的闭环系统辨识新方法

徐宝昌[†], 林忠华, 肖玉月

(中国石油大学(北京) 自动化系, 北京 102249)

摘要: 本文基于近似最小一乘准则和主成分分析, 针对反馈通道模型阶次低于前向通道模型阶次且反馈通道不存在噪声的闭环系统, 进行了近似偏最小一乘递推辨识算法的推导. 为解决最小一乘准则函数不可微的问题, 本文算法用确定性可导函数近似代替残差绝对值. 近似偏最小一乘辨识算法可以克服基于最小二乘准则的辨识算法在受到满足(S α S)分布的尖峰噪声干扰时残差平方项过大的缺点, 具有目标函数可导, 计算简单的优点. 同时, 通过主成分分析去除数据向量各元素之间的线性相关, 可以得出模型参数的唯一解. 仿真实验表明, 本文算法可以对反馈通道模型阶次低于前向通道模型阶次的闭环系统进行直接辨识, 抑制了尖峰噪声对辨识结果的影响, 具有优良的稳健性, 可以更好地应用于闭环系统辨识.

关键词: 偏最小一乘; 闭环系统; 主成分分析; 相关性; 尖峰噪声

中图分类号: TP13 **文献标识码:** A

Identification of closed-loop system by partial least absolute deviation

XU Bao-chang[†], LIN Zhong-hua, XIAO Yu-yue

(Department of Automation, China University of Petroleum (Beijing), Beijing 102249, China)

Abstract: Based on approximate least absolute deviation criterion and principal component analysis, a recursive partial approximate least absolute deviation (PALAD) identification algorithm is deduced for closed-loop system whose model order of feedback channel is lower than that of the forward channel and there is no noise in the feedback channel. To solve the non-differentiable problem of the least absolute deviation, a deterministic derivable function is established to approximate the absolute value under certain situations in this paper. The proposed method can overcome the disadvantage of large square residual of least square criterion when the identification data is disturbed by the impulse noise which obeys symmetrical alpha stable distribution(S α S). By adopting principal component analysis to eliminate the linear correlation among the elements of data vector, the unique solution of model parameters can be easily acquired by the proposed method. The simulation experiments show that the proposed method can be directly used to identify closed-loop system whose model order of feedback channel is lower than that of the forward channel. Moreover, the proposed algorithm can restrain the impact of impulse noise effectively, has strong robustness and can be better applied to closed-loop identification..

Key words: partial least absolute deviation; closed-loop systems; principal component analysis; correlation; impulse noise

1 引言(Introduction)

系统辨识从进行辨识的实验条件来说, 可以分为开环辨识和闭环辨识两大类. 对于开环系统, 已有许多成熟的经典辨识方法, 如相关分析法、辅助变量法、频谱分析法和预报误差法等, 但这些方法不能直接应用于闭环条件下对象的辨识^[1-3]. 然而, 一些实际运行着的工业过程, 不能轻易切断其中的反馈回路, 否则可能会造成过程失控, 对系统运行安全性、经济性产生严重影响, 因此, 对闭环辨识的研究尤为重要. 对于闭环系统, 存在3种经典的辨识方法: 直接法、间

接法、联合输入输出法^[4-8]. 直接法在一定条件下利用闭环系统输入输出数据, 把对象当作在开环条件下进行辨识, 间接法与联合输入输出法实验前则要知道反馈控制器的传递函数, 并且需要至少3个可检测的信号^[9], 直接法比间接法和联合输入输出法更简洁. 文献[10]的研究表明, 当反馈通道的模型阶次低于前向通道的模型阶次且反馈通道不存在干扰噪声时, 由于数据向量中的各个元素之间存在线性相关, 导致直接辨识算法的解不唯一, 无法估计模型的参数.

除数据向量的线性相关外, 随机噪声也是影响闭

收稿日期: 2015-10-31; 录用日期: 2016-09-22.

[†]通信作者. E-mail: xbcyl@163.com; Tel.: +86 13683262121.

本人责任编辑: 李少远.

国家重大科技专项(2011ZX05021-003)资助.

Support by Important National Science & Technology Specific Projects of China (2011ZX05021-003).

环系统辨识结果的重要因素. 当随机噪声满足正态分布(为白噪声)时, 基于最小二乘准则的辨识算法具有优良的性质^[11]. 基于最小二乘准则的偏最小二乘方法(partial least squares, PLS)广泛应用于统计数据处理、软测量建模和过程监控等领域^[12-15], 采用主成分分析去除自变量之间的相关性, 解决了采用普通最小二乘方法估计参数出现病态解的问题. 另外, 也有学者将偏最小二乘方法运用于动态系统辨识算法研究: 尹力等^[16]提出一种基于偏最小二乘回归的铣削力多项式模型系数辨识方法, 具有实验次数少、计算简单的特点; Antoine等^[17]将偏最小二乘方法应用于自回归滑动平均模型(vector autoregressive moving average, VARMA)辨识, 并与基于平衡实现理论和基于典型变量分析的辨识算法进行对比; Lauri等^[18]提出了一种基于PLS的带控制量的自回归(autoregressive exogenous, ARX)预测模型辨识方法. 上述研究主要针对开环系统模型参数辨识, 没有考虑利用PLS解决闭环系统辨识的相关性问题. 此外, 当随机噪声不满足正态分布时, 如系统存在满足 $S\alpha S$ 分布(对称 α 稳定分布)的尖峰噪声, 由于其目标函数中的平方项, 数据存在很小的改变都会引起较大的波动, 对辨识结果产生很大的影响^[19-20], 算法稳健性差. 而最小一乘准则将残差绝对值和作为目标函数, 由于只对误差作一次方运算, 辨识算法对尖峰噪声的敏感度大大降低, 增强了算法稳健性. 此时, 最小一乘估计的统计性能优于最小二乘估计^[21], 其具有不可替代的优越性. 然而, 最小一乘准则函数不可微, 需要求解一个非光滑的优化问题^[22-23], 给计算增加了很大难度.

为解决闭环系统辨识的相关性问题, 本文通过主成分分析, 将存在线性相关的闭环辨识观测数据向量投影到非线性相关的低维特征空间中, 去除相关性, 为开环辨识算法应用到闭环系统辨识创造条件. 针对最小一乘准则不可微的缺点, 本文引进一个可导确定性函数近似等价绝对值函数, 构造近似最小一乘准则目标函数, 在此基础上推导了基于近似最小一乘准则的偏最小一乘闭环系统辨识算法(partial approximate least absolute deviation, PALAD). 本文算法在保持最小一乘准则算法稳健性的同时, 具有目标函数可导、计算简单的优点. 仿真实验表明, 本文算法可以去除闭环系统数据向量的线性相关, 得到闭环系统模型参数的唯一解, 可以有效克服尖峰噪声对闭环系统参数辨识的影响, 与基于最小二乘准则的辨识算法进行比较分析, 其稳健性更强, 辨识精度更高, 充分体现了本文算法的优越性.

2 闭环系统模型描述(Model formulation of closed-loop system)

考虑闭环系统模型如图1所示, 由一条前向通道和一条反馈通道组成.

图1中: $G_s(z^{-1})$ 为系统前向通道的传递函数;

$G_c(z^{-1})$ 为反馈通道的控制器传递函数; $N_p(z^{-1})$ 和 $N_\zeta(z^{-1})$ 分别为前向通道噪声 $v(k)$ 和反馈通道噪声 $\zeta(k)$ 的滤波器; $N_p(z^{-1})$ 为扰动信号 $p(k)$ 的滤波器; z^{-1} 是后移算子, $z^{-1}v(k) = v(k-1)$; $v(k)$ 和 $\zeta(k)$ 都是均值为0、方差为 σ_v^2 和 σ_ζ^2 的互不相关平稳随机噪声; $p(k)$ 是可测的; $r(k)$ 为给定值信号; $y(k)$ 为输出信号.

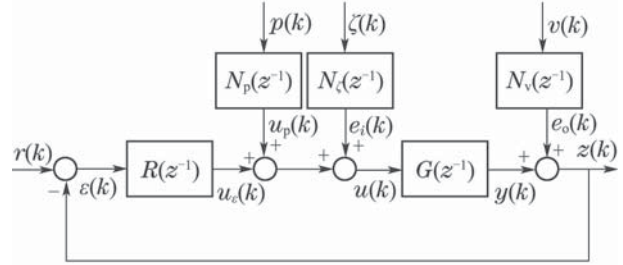


图1 闭环系统模型结构

Fig. 1 Model structure of closed-loop system

系统中传递函数和滤波器的模型结构取为

$$\begin{aligned} G_s(z^{-1}) &= \frac{B(z^{-1})}{A(z^{-1})} z^{-d}, \quad G_c(z^{-1}) = \frac{Q(z^{-1})}{P(z^{-1})}, \\ N_v(z^{-1}) &= \frac{D(z^{-1})}{A(z^{-1})}, \quad N_\zeta(z^{-1}) = \frac{F(z^{-1})}{P(z^{-1})}, \\ N_p(z^{-1}) &= \frac{E(z^{-1})}{P(z^{-1})}, \end{aligned}$$

其中:

$$\begin{aligned} A(z^{-1}) &= 1 + a_1 z^{-1} + \cdots + a_{n_a} z^{-n_a}, \\ B(z^{-1}) &= b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b}, \\ D(z^{-1}) &= 1 + d_1 z^{-1} + \cdots + d_{n_d} z^{-n_d}, \\ P(z^{-1}) &= 1 + p_1 z^{-1} + \cdots + p_{n_p} z^{-n_p}, \\ Q(z^{-1}) &= q_0 + q_1 z^{-1} + \cdots + q_{n_q} z^{-n_q}, \\ F(z^{-1}) &= 1 + f_1 z^{-1} + \cdots + f_{n_f} z^{-n_f}, \\ E(z^{-1}) &= 1 + e_1 z^{-1} + \cdots + e_{n_e} z^{-n_e}. \end{aligned}$$

考虑给定值信号 $r(k) = 0$, 扰动信号 $p(k) = 0$, 纯滞后 $d = 0$, 前向通道噪声滤波器 $D(z^{-1}) = 1$, 即前向通道只受到白噪声 $v(k)$ 干扰, 系统模型的待辨识参数为

$$\theta = [a_1 \ a_2 \ \cdots \ a_{n_a} \ b_1 \ b_2 \ \cdots \ b_{n_b}]^T. \quad (1)$$

令

$$\varphi(k) = [-y(k-1) \ -y(k-2) \ \cdots \ -y(k-n_a) \ u(k-1) \ u(k-2) \ \cdots \ u(k-n_b)]^T. \quad (2)$$

$$\text{令 } P(z^{-1}) = 1, \quad G_c(z^{-1}) = \frac{Q(z^{-1})}{P(z^{-1})} = q_0 + q_1 z^{-1} + \cdots + q_{n_q} z^{-n_q}.$$

则闭环系统可以表示为

$$\begin{cases} y(k) = \varphi^T(k) \theta + v(k), \\ u(k) = G_c(z^{-1}) y(k). \end{cases} \quad (3)$$

当反馈通道上不存在噪声干扰时, 即 $\zeta(k) = 0$, 若系统反馈通道的模型阶次低于前向通道的模型阶次, 由于数据向量 $\varphi(k)$ 中的元素 $u(k-1)$ 通过反馈通道后与元素 $y(k-1), y(k-2), \dots, y(k-n_a)$ 线性相关, 模型参数解不唯一, 无法得到该类闭环系统的参数估计值. 本文算法可去除 $\varphi(k)$ 中元素的相关性, 通过直接辨识算法得到闭环系统模型参数的唯一解.

3 基于近似偏最小一乘的闭环系统辨识方法 (Identification method of closed-loop system based on partial approximate least absolute deviation)

本文算法通过主成分分析得到表征数据矩阵 $\Phi = [\varphi(1) \ \varphi(2) \ \dots \ \varphi(l)]^T$ 的特征向量 ω 、主成分向量 t , 以及包含 Φ 的信息的负荷矩阵 Ψ , 通过分析 Φ 和 Ψ 之间的关系, 将存在线性相关的 Φ 转换为满秩的 Ψ , 为求解数据矩阵存在线性相关的闭环系统模型参数的唯一解奠定基础. 在主成分分析的基础上, 对闭环系统输入输出数据进行偏最小一乘回归, 得到闭环系统参数估计值, 由于最小一乘准则函数不可微, 为减小计算量, 引入一个确定性可导函数近似代替残差绝对值. 本文接下来首先介绍主成分分析的方法、分析 Φ 和 Ψ 之间的关系, 然后用一个确定的函数逼近于绝对值误差准则, 解决绝对值函数不可微的缺点, 最后推导基于主成分分析的近似偏最小一乘回归递推辨识算法.

3.1 主成分分析 (Principal component analysis)

主成分分析通过将存在线性相关的数据矩阵投影到无线性相关的低维特征空间, 得到数据矩阵 Φ 的相互正交的特征向量, 进而得到蕴含数据矩阵 Φ 信息的负荷矩阵 Ψ . 为使主成分在最大限度概括原数据矩阵的数据变化信息的同时, 尽可能多地考虑数据矩阵主成分对于输出数据变化的解释作用, 采用如下准则函数进行主成分提取^[24]:

$$J = \min_{\|\omega=1\|} \sum_{k=1}^n [\|\varphi(k) - \omega \omega^T \varphi(k)\|^2 + \|y(k) - \omega^T \varphi(k)\|^2], \quad (4)$$

式中: $\varphi(k)$ 为数据向量, $y(k)$ 为输出数据, ω 为主成分权值系数向量, n 为样本数.

主成分分析具体步骤如下:

1) 取输出数据向量 y 作为主成分向量 x 的初始值: $x = y$.

2) 计算数据矩阵 Φ 的权值向量 ω :

$$\omega = x^T \Phi / (x^T x). \quad (5)$$

3) 归一化权值向量 ω :

$$\omega_{\text{new}}^T = \omega^T / \|\omega^T\|. \quad (6)$$

4) 计算 Φ 的主成分向量:

$$t = \Phi \omega_{\text{new}} / (\omega_{\text{new}}^T \omega_{\text{new}}). \quad (7)$$

5) 计算 y 的权值向量:

$$c^T = t^T y / (t^T t). \quad (8)$$

6) 归一化权值向量 c :

$$c_{\text{new}}^T = c^T / \|c^T\|. \quad (9)$$

7) 计算 y 的主成分向量:

$$x = y c_{\text{new}} / (c_{\text{new}}^T c_{\text{new}}). \quad (10)$$

若第4)步得出的主成分向量 t_i^k 与上次迭代得出的 t_i^{k-1} 相等或者其差在预定的允许误差范围内, 则说明 t_i^k 已收敛于其真值, 此时可执行第8)步, 以进行下一个主成分 t_{i+1} 的计算; 否则转至第2)步, 对 ω, c, t, x 进行第 $k+1$ 次迭代.

8) 计算 Φ 相对于 t 的负荷向量:

$$\psi^T = t^T \Phi / (t^T t). \quad (11)$$

9) 计算 y 相对于 x 的负荷向量:

$$h^T = x^T y / (x^T x). \quad (12)$$

10) 计算 x 对 t 的回归系数:

$$\beta = x^T t / (t^T t) = t^T x / (t^T t). \quad (13)$$

11) 计算残差矩阵 K 和残差向量 f :

$$\begin{cases} K = \Phi - t \psi^T, \\ f = y - \beta t h^T. \end{cases} \quad (14)$$

令 $\Phi = K, y = f$, 重复以上步骤, 计算下一个主成分向量, 直至 Φ 的秩为0, 全部主成分计算完毕, 得到 Φ 的负荷矩阵 $\Psi = [\psi_1 \ \psi_2 \ \dots \ \psi_s]$, 其中 $s = \text{rank}(\Phi)$. 数据矩阵 Φ 与负荷矩阵 Ψ 满足以下定理^[25]:

定理 1 假设 Φ 为 $n \times m$ ($m = n_a + n_b$) 矩阵, 如果 $\text{rank}(\Phi) = s \leq m$, 在提取了第 s 个主成分 t_s 后, 数据矩阵的残差矩阵 $K_s = 0 \in \mathbb{R}^{n \times s}$. 特别定义 t_i 为归一化的特征向量, 即 $T^T T = I$, 则有

$$\Phi^T \Phi = (T_s \Psi^T + K_s)^T \Psi^T + K_s = \Psi \Psi^T,$$

其中:

$$T = [t_1 \ t_2 \ \dots \ t_s \ t_{s+1} \ \dots \ t_m] \in \mathbb{R}^{n \times m},$$

$$T_s = [t_1 \ t_2 \ \dots \ t_s] \in \mathbb{R}^{n \times s},$$

$$\Psi = [\psi_1 \ \psi_2 \ \dots \ \psi_s] \in \mathbb{R}^{m \times s}.$$

3.2 近似最小一乘准则 (Approximate least absolute deviation criterion)

为解决绝对值函数不可微的缺点, 用一个确定的函数逼近于绝对值误差准则 (L1-范数), 下面的对数函数可以满足这个要求:

$$J(w) = \eta \ln \left[\cosh \frac{w}{\eta} \right]. \quad (15)$$

式(15)中: η 是一个可控参数, 非线性函数 $J(w)$ 与 η 相关. 当 η 足够小的时候, 非线性函数 $\eta \ln \left[\cosh \frac{w}{\eta} \right] \approx |w|$, 即 $J(w) \approx |w|$, 逼近于绝对值误差准则函数. 取 $\eta = 0.01$, 该函数可以很好地近似绝对值函数. 对于可以用式(3)描述的闭环系统, 为了提高辨识结果的准确性和稳健性, 采用式(16)的近似最小一乘函数作为辨识的准则函数:

$$J(\theta) = \sum_{k=1}^m \eta \ln \cosh \left[\frac{y(k) - \varphi^T(k)\theta}{\eta} \right], \quad (16)$$

当 η 很小时, 此准则函数可以很好地近似绝对值函数, 这样就解决了非光滑函数优化的难题.

3.3 算法推导(Algorithm derivation)

对于由式(3)描述的闭环系统, 其最小一乘目标函数可近似表示为式(16), 利用Gauss-Newton法求解目标函数的最小化问题, 可得系统参数的递推计算公式为

$$\theta(m) = \theta(m-1) + \tau \Delta \theta(m), \quad (17)$$

其中: τ 为学习速率, $\Delta \theta(m)$ 为Gauss-Newton的搜索方向, 定义为

$$\Delta \theta(m) = -[H_m]^{-1} g(m), \quad (18)$$

式中: g 是 $J(\theta)$ 关于 θ 的梯度, H 是 $J(\theta)$ 的Hessian矩阵.

由于

$$J_m = J_{m-1} + \eta \ln \left[\cosh \frac{e(m)}{\eta} \right], \quad (19)$$

令 $e(m) = y(m) - \varphi^T(m)\theta(m-1)$, 即有

$$J_m = J_{m-1} + \eta \ln \left[\cosh \frac{e(m)}{\eta} \right], \quad (20)$$

所以

$$\begin{aligned} g(m) &= \frac{\partial J_m}{\partial \theta} \Big|_{\theta=\theta(m-1)} = \\ &= \frac{\partial J_{m-1}}{\partial \theta} \Big|_{\theta=\theta(m-1)} - \varphi(m) \tanh \left[\frac{e(m)}{\eta} \right] = \\ &= g(m-1) - \varphi(m) \tanh \left[\frac{e(m)}{\eta} \right], \end{aligned} \quad (21)$$

$$\begin{aligned} H_m &= \frac{\partial^2 J_m}{\partial \theta^2} = \\ &= \frac{1}{\eta} \sum_{k=1}^m \varphi(k) \varphi^T(k) \left[1 - \tanh^2 \frac{e(k)}{\eta} \right]. \end{aligned} \quad (22)$$

显然, Hessian矩阵是对称矩阵, 为保证算法的稳定性, 在递推计算过程中必须要保证它的正定性, 以使搜索的方向始终是指向“下山”的方向.

$$H_m = \frac{\partial^2 J_m}{\partial \theta^2} = H_{m-1} + \frac{1}{\eta} \varphi(m) \varphi^T(m) \left[1 - \right.$$

$$\left. \tanh^2 \frac{e(m)}{\eta} \right], \quad (23)$$

令 $\lambda(m) = \frac{1}{\eta} (1 - \tanh^2 \frac{e(m)}{\eta})$, 则

$$\begin{aligned} H_m &= \frac{\partial^2 J_m}{\partial \theta^2} = H_{m-1} + \lambda(m) \varphi(m) \varphi^T(m) = \\ &= H_{m-1} + \sqrt{\lambda(m)} \varphi(m) (\sqrt{\lambda(m)} \varphi(m))^T. \end{aligned} \quad (24)$$

根据定理1, 数据矩阵 $\Phi(m)$ 与其负荷矩阵 $\Psi(m)$ 满足以下关系:

$$\begin{aligned} \Psi(m) \Psi^T(m) &= \Phi^T(m) \Phi(m) = \\ &= \Phi^T(m-1) \Phi(m-1) + \varphi(m) \varphi^T(m) = \\ &= \Psi(m-1) \Psi^T(m-1) + \varphi(m) \varphi^T(m), \end{aligned} \quad (25)$$

$$\text{其中 } \Phi(m) = \begin{bmatrix} \Phi(m-1) \\ \varphi^T(m) \end{bmatrix}.$$

比较式(24)和式(25), 可以对 H_m 作如下处理:

$$\begin{aligned} H_m &= \bar{\Psi}(m) \bar{\Psi}^T(m) = \\ &= \bar{\Psi}(m-1) \bar{\Psi}^T(m-1) + \\ &= \sqrt{\lambda(m)} \varphi(m) (\sqrt{\lambda(m)} \varphi(m))^T = \\ &= \bar{\Psi}(m-1) \bar{\Psi}^T(m-1) + \bar{\varphi}(m) \bar{\varphi}^T(m). \end{aligned} \quad (26)$$

式(21)中: $\bar{\varphi}(m) = \sqrt{\lambda(m)} \varphi(m)$, $\bar{\Psi}(m)$ 为 $\sqrt{\lambda(m)} \Phi(m)$ 的负荷矩阵, 可通过主成分分析方法计算.

$$\begin{aligned} \Delta \theta(m) &= -[H_m]^{-1} g(m) = \\ &= -[\bar{\Psi}(m-1) \bar{\Psi}^T(m-1) + \bar{\varphi}(m) \bar{\varphi}^T(m)]^{-1} \cdot \\ &= \{g(m-1) - \varphi(m) \tanh \left[\frac{e(m)}{\eta} \right]\}. \end{aligned} \quad (27)$$

由式(17)知道 $\theta(m) = \theta(m-1) + \tau \Delta \theta(m)$.

综上, 可以得到近似偏最小一乘算法的递推公式. 其中, τ 为学习速率, $\theta(0)$ 和 $\Delta \theta(0)$ 均初始化为充分小的实向量, 由随机数组成. 每次增加一组测试数据, 系统参数 θ 就进行一次修正, 直至整个辨识过程结束.

4 仿真实验(Simulations)

当闭环系统的反馈通道模型阶次低于前向通道模型阶次时, 只有在反馈通道存在噪声干扰时, 模型的参数估计值才是唯一的、一致的, 如果反馈通道不存在噪声, 则因为观测数据矩阵中各变量之间存在线性相关, 从而导致以往的辨识算法不能得到模型参数的唯一解. PALAD通过主成分提取, 去除观测数据矩阵中的相关性, 从而可以在反馈通道模型阶次低于前向通道模型阶次且反馈通道不存在噪声时, 有效地辨识出前向通道的模型参数. 本文针对反馈通道模型阶次低于前向通道模型阶次且反馈通道不存在噪声扰动的情况, 在白噪声和尖峰噪声等不同类型噪声干扰下进行仿真实验, 分析了PALAD算法的辨识性能, 并与偏最小二乘算法(partial least squares, PLS)、最小一乘

算法(least absolute deviation, LAD)和偏最小一乘算法(partial least absolute deviation, PLAD)进行对比.

考虑闭环系统前向通道模型为

$$y(k) = -1.4y(k-1) - 0.2y(k-2) + 0.35u(k-1) + 0.11u(k-2) + v(k).$$

反馈通道模型为

$$u(k) = 0.2y(k) - 3y(k-1),$$

其中: 输入序列 $u(k)$ 为均值为零、方差为 $\sigma_u^2 = 1$ 的持续激励随机序列; 噪声序列 $v(k)$ 为均值为零、方差 $\sigma_v^2 = 1$ 的随机白噪声序列. 利用模型产生输出观测数据, 根据本文算法用MATLAB编程, 最终实现系统参数向量的估计. 为保证仿真实验的可靠性, 对每一种仿真实验都进行了100次数据长度 $L = 3000$ 的重复独立试验, 以相对误差 δ 和收敛速度作为评价标准.

仿真实验中的尖峰噪声满足 $S\alpha S$ 分布^[26], 为对称的 α 稳定分布, 标准 $S\alpha S$ 分布的概率密度函数为

$$f_{\alpha}(x) = \begin{cases} \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} \Gamma(\alpha k + 1) |x|^{-\alpha k} \dots \\ \sin(k\alpha\pi/2), 0 < \alpha < 1; \\ \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{(-1)^k}{2k!} \Gamma[(2k+1)/\alpha] x^{2k}, \\ 1 < \alpha \leq 2, \end{cases}$$

表 1 前向通道仅有白噪声干扰时的辨识结果

Table 1 The identification results when forward channel is disturbed by white noise only

算 法	a_1	a_2	b_1	b_2	$\delta/\%$	步数/N	t/s
PALAD	-1.4384	-0.2125	0.3499	0.1113	2.3405	405	0.9236
PLS	-1.3832	-0.2068	0.3420	0.1026	1.4875	190	0.7315
LAD	-1.3324	-1.0767	0.0509	0.1015	63.5722	—	2.5728
PLAD	-1.3779	-0.2055	0.3409	0.1017	1.6815	—	2.3763
真值	-1.4000	-0.2000	0.3500	0.1100	0.0000	—	—

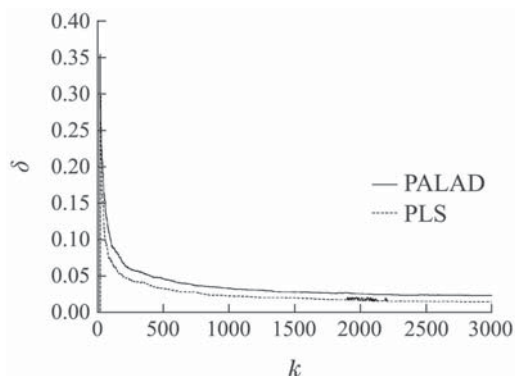


图 2 仅存在白噪声干扰时的相对误差曲线

Fig. 2 Relative error curves when forward channel is disturbed by white noise only

表1中, a_1, a_2, b_1, b_2 为模型参数. 从表1可以看出, 当反馈通道模型阶次低于前向通道模型阶次且

式中: $\Gamma(\cdot)$ 为伽马函数, α 为特征指数, α 的取值越小, 满足 $S\alpha S$ 分布的随机变量大幅值样本发生的概率越大, 随机变量样本值的脉冲性就越强.

标准 $S\alpha S$ 分布随机变量可以通过以下方法获得^[27]: 若 V 是在 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 上满足均匀分布的随机变量, W 是一个均值为1的指数分布的随机变量, 并且 V 和 W 相互独立, 则

$$Y = \frac{\sin(\alpha V)}{(\cos V)^{1/\alpha}} \left(\frac{\cos((1-\alpha)V)}{W} \right)^{(1-\alpha)/\alpha}$$

服从标准 $S\alpha S$ 分布.

4.1 前向通道仅存在白噪声干扰(Forward channel with white noise only)

当前向通道仅有白噪声干扰时, PALAD 算法、PLS算法、LAD算法和PLAD算法的辨识结果如表1所示, 其中, 步数 N 为 $\delta < 5\%$ 所需要的样本数, 时间为算法运行时间. 由于绝对值函数不可微, 本文仿真中的LAD算法采用线性规划求解模型参数估计值, PLAD算法中主成分相对于输出数据的最小一乘回归系数也通过线性规划求解, 线性规划算法需要采集所有的样本之后才能够进行优化, 其搜索策略与PALAD算法和PLS算法不同^[28], 因此, 仅比较PALAD算法和PLS算法递推过程的相对误差曲线, 如图2所示.

反馈通道不存在噪声时, LAD算法由于其数据矩阵存在线性相关, 其参数估计值的解不稳定, 各次仿真实验的参数估计值相差较大且辨识精度差, 无法得到闭环系统的模型参数估计值, 因此, 在本文的其他仿真实验中, 不再对比LAD算法的辨识结果. PALAD 算法、PLAD算法和PLS算法均能对前向通道模型参数进行直接辨识, 克服了以往辨识算法在此条件下不能得到闭环系统模型参数唯一解的缺点, 但PLAD算法的计算时间显著长于PALAD 算法和 PLS 算法. 结合图2, PALAD算法和PLS算法的辨识精度和收敛速度均较为理想, 但是二者相比较之下, PLS算法的相对误差下降至5%所需的步数 N 明显低于PALAD 算法, 且前者的辨识精度更高. 综上, 在前向通道仅存在白噪声干扰时, PLS法较基于最小一乘准则的LAD算法、PALAD算法

和PLAD 算法具有更优良的辨识性能.

4.2 前向通道仅存在尖峰噪声干扰 (Forward channel with impulse noise only)

当前向通道仅存在尖峰噪声干扰时, 尖峰噪声此时作为闭环系统的激励信号, 系统在大幅值信号的激励条件下, 辨识前向通道的模型参数. 为充分

分析尖峰噪声对辨识结果的影响, 本文分别采用不同强度的尖峰噪声激励系统, 即分别采用特征指数为 $\alpha = 1.5$, $\alpha = 1.2$ 和 $\alpha = 0.9$ 的尖峰噪声, α 的值越小, 噪声信号出现大幅值样本的概率越高. PALAD算法、PLS算法和PLAD算法的辨识结果如表2~4所示, 相对误差曲线如图3~5所示.

表 2 当 $\alpha = 1.5$ 时的辨识结果
Table 2 The identification results when $\alpha = 1.5$

算 法	a_1	a_2	b_1	b_2	$\delta / \%$	步数 / N	t / s
PALAD	-1.4115	-0.2093	0.3504	0.1134	1.0460	174	0.9242
PLS	-1.4153	-0.2119	0.3529	0.1040	1.4033	122	0.7322
PLAD	-1.3967	-0.2087	0.3467	0.1097	0.6741	—	2.4815
真值	-1.4000	-0.2000	0.3500	0.1100	0.0000	—	—

表 3 当 $\alpha = 1.2$ 时的辨识结果
Table 3 The identification results when $\alpha = 1.2$

算 法	a_1	a_2	b_1	b_2	$\delta / \%$	步数 / N	t / s
PALAD	-1.3875	-0.2062	0.3425	0.1061	1.1273	100	0.9301
PLS	-1.3821	-0.2057	0.3408	0.1023	1.5322	95	0.7501
PLAD	-1.4038	-0.2092	0.3467	0.1088	0.7198	—	2.5458
真值	-1.4000	-0.2000	0.3500	0.1100	0.0000	—	—

表 4 当 $\alpha = 0.9$ 时的辨识结果
Table 4 The identification results when $\alpha = 0.9$

算 法	a_1	a_2	b_1	b_2	$\delta / \%$	步数 / N	t / s
PALAD	-1.4206	-0.2145	0.3525	0.1196	1.3156	57	0.8908
PLS	-1.3832	-0.2068	0.3420	0.1026	1.3950	88	0.7229
PLAD	-1.3900	-0.2073	0.3438	0.1070	0.9690	—	2.7507
真值	-1.4000	-0.2000	0.3500	0.1100	0.0000	—	—

从表2~4可以看出, 当前向通道仅受到尖峰噪声干扰时, PALAD算法、PLS算法和PLAD算法都有较高的辨识精度, 在不同强度尖峰噪声干扰下, 辨识精度变化不大. 其中, PLAD算法的辨识精度最高, 但是该算法采用线性规划求解回归系数最优值, 不能应用于在线辨识. 另一方面, 随着尖峰噪声特征指数减小, PALAD算法和PLS算法的相对误差下降至5%所需的步数逐渐减少, PALAD算法的减幅达到43%, 高于PLS算法. 当 $\alpha = 1.5$ 时, $N_{\text{PALAD}} > N_{\text{PLS}}$, PLS算法收敛速度显著高于PALAD算法; 当 $\alpha = 1.2$ 时, N_{PALAD} 与 N_{PLS} 相差不大, PLS算法收敛速度略快于PALAD算法; 当 $\alpha = 0.9$ 时, $N_{\text{PALAD}} < N_{\text{PLS}}$, PLS算法收敛速度明显低于PALAD算法. 随着尖峰噪声强度的增加, PALAD算法的收敛速度快速提高, 并逐渐超越PLS算法.

从图3~5可以看出, 在不同强度尖峰噪声干扰下, PALAD算法的辨识精度高于PLS算法. 同时注意到, PLS算法的相对误差曲线有波动现象, 特征指数 α 的

值越小, PLS相对误差曲线波动越剧烈, 而PALAD相对误差曲线更加平稳, 因此, 在尖峰噪声干扰时, PALAD算法的辨识结果具有更高的稳健性. 综上所述, 在前向通道仅存在尖峰噪声干扰的情况下, PALAD算法具有更加稳健的辨识结果, 收敛速度较快, 并且随着 α 的减小, PALAD算法的优势更加显著.

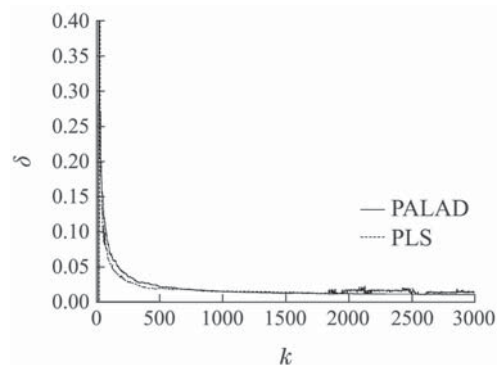
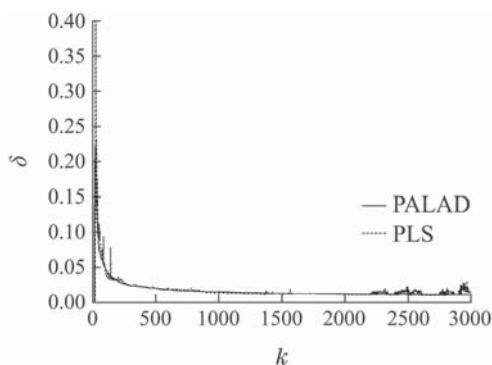
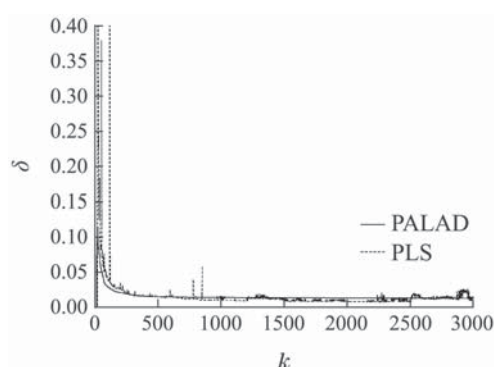


图 3 当 $\alpha = 1.5$ 时的相对误差曲线
Fig. 3 Relative error curves when $\alpha = 1.5$

图 4 当 $\alpha = 1.2$ 时的相对误差曲线Fig. 4 Relative error curves when $\alpha = 1.2$ 图 5 当 $\alpha = 0.9$ 时的相对误差曲线Fig. 5 Relative error curves when $\alpha = 0.9$

4.3 前向通道同时存在白噪声和尖峰噪声干扰 (Forward channel with white noise and impulse noise)

当前向通道同时存在白噪声和尖峰噪声干扰时,

取白噪声的方差为 $\sigma_v^2 = 1$, 尖峰噪声特征指数 $\alpha = 0.9$ 进行仿真实验, PALAD算法、PLS算法和PLAD算法的辨识结果, 如表5所示, 相对误差曲线, 如图6所示。

表 5 存在白噪声和尖峰噪声的辨识结果

Table 5 The identification results when forward channel is disturbed by white noise and impulse noise

算 法	a_1	a_2	b_1	b_2	$\delta/\%$	步数/N	t/s
PALAD	-1.4206	-0.2139	0.3525	0.1196	1.8391	62	0.9149
PLS	-1.3827	-0.2184	0.3398	0.1005	1.9677	113	0.7357
PLAD	-1.3831	-0.2064	0.3426	0.1031	1.4201	—	2.5575
真值	-1.4000	-0.2000	0.3500	0.1100	0.0000	—	—

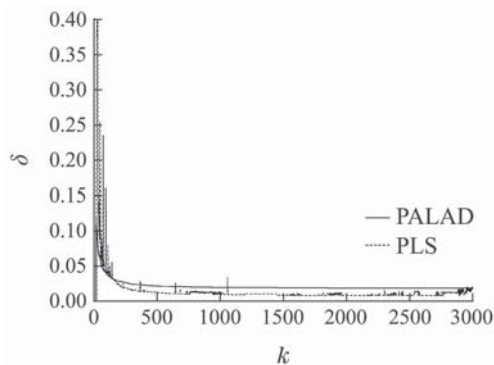


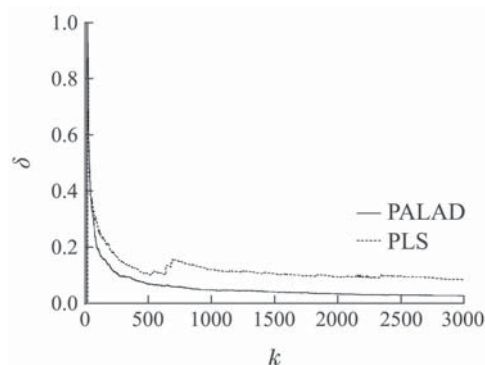
图 6 存在白噪声和尖峰噪声时的相对误差曲线

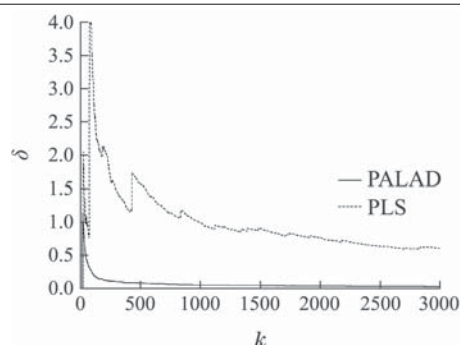
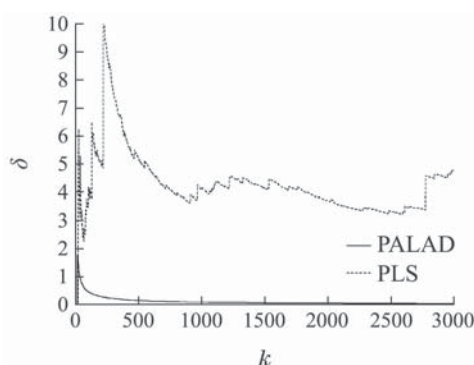
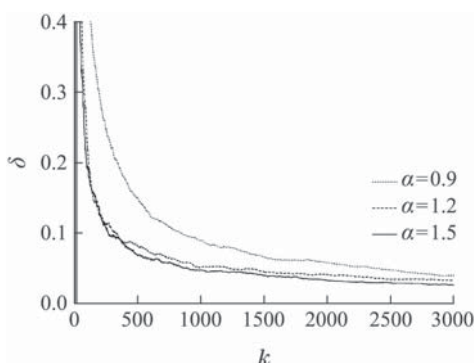
Fig. 6 Relative error curves when forward channel is disturbed by white noise and impulse noise

结合表4-5和图6, 在白噪声和尖峰噪声同时存在的情况下, PALAD算法和PLAD算法的辨识精度高于仅存在白噪声的情况, 低于仅存在尖峰噪声的情况; PLS算法的相对误差曲线存在较大波动, 参数估计值的可靠性低于PALAD算法. 同时, PALAD算法的相对误差小于5%所需样本数明显少于PLS算法, 有效减少辨识时间。

文献[19-20]的研究表明, 在开环系统中, 尖峰噪声对基于最小二乘准则的辨识算法有很大影响,

PLS在尖峰噪声干扰下, 无法有效地辨识模型参数. 而在闭环系统中, 尖峰噪声通过反馈通道以后, 作为系统的输入作用于前向通道, 此过程具有削弱尖峰噪声干扰的作用, 因此, PLS算法能够得到较为准确的参数估计值. 为进一步比较PALAD算法和PLS算法的辨识性能, 考虑当白噪声和尖峰噪声干扰同时存在时, 尖峰噪声只影响输出信号, 而不进入反馈通道, 影响下一时刻的模型输出. 此种情况下, 分别取 $\alpha = 1.5$, $\alpha = 1.2$ 和 $\alpha = 0.9$ 进行仿真, PALAD算法和PLS算法的相对误差曲线如图7所示。

图 7(a) 当 $\alpha = 1.5$ 时的辨识结果Fig. 7(a) Relative error curves when $\alpha = 1.5$

图 7(b) 当 $\alpha = 1.2$ 时的辨识结果Fig. 7(b) Relative error curves when $\alpha = 1.2$ 图 7(c) 当 $\alpha = 0.9$ 时的辨识结果Fig. 7(c) Relative error curves when $\alpha = 0.9$ 图 7(d) PALAD 算法在不同 α 时的相对误差曲线Fig. 7(d) Relative error curves of PALAD with different α

由图7(a)–7(c)可知, 当 $\alpha = 1.5$ 时, 尖峰噪声的干扰强度较小, 由于受到尖峰噪声的影响, PLS 算法的相对误差曲线产生波动, 与图3相比较, PLS 算法的辨识精度大幅下降; 随着 α 的减小, 尖峰噪声干扰强度增大, PLS 算法的相对误差曲线发生剧烈震荡, 无法进行有效的模型参数辨识. 与之相反, PALAD 算法在不同强度的尖峰噪声干扰下, 均能得到较高的辨识精度, 参数估计值接近真实值, 从图7(d)可以看出, 随着 α 的减小, PALAD 算法的相对误差有所增加, 但变化幅度不大, 体现了 PALAD 算法的稳健性. 因此, 在尖峰噪声不进入反馈通道的情况下, PALAD 算法可以有效抑制大幅值噪声干扰, 具有更

强的稳健性.

4.4 辨识结果分析 (Analysis of identification results)

分析以上仿真结果, 可以得到下列结论:

1) 本文算法可以有效地辨识出闭环系统的模型参数;

2) 在闭环系统反馈通道模型阶次低于前向通道模型阶次时, 本文算法通过提取主成分, 去除观测矩阵各向量之间的相关性, 对闭环系统进行直接辨识, 可以得到闭环系统模型参数估计值的唯一解;

3) 在不存在尖峰噪声干扰的情况下, PLS 算法的辨识精度略高, 收敛速度更快, 辨识性能优于 PALAD 算法和 PLAD 算法;

4) 当测量数据中存在尖峰噪声时, PLS 算法的相对误差曲线发生波动, 辨识精度不稳定, 随着特征指数的减小, 波动越剧烈; PALAD 算法和 PLAD 算法有效抑制尖峰噪声的干扰作用, 在不同强度尖峰噪声的干扰下, 相对误差保持平稳, 并收敛至较小值, 随着特征指数的减小, PALAD 算法的精度有所降低, 但其相对误差变化不大, 这体现了基于最小一乘准则辨识算法良好的稳健性; PLAD 算法局限于离线参数辨识, PALAD 算法可在线辨识模型参数, 具有更广阔的应用空间.

5 结论 (Conclusions)

针对反馈通道的模型阶次低于前向通道的模型阶次且反馈通道无噪声干扰的闭环系统, 基于近似最小一乘准则, 结合主成分分析, 推导了基于近似偏最小一乘的闭环系统辨识算法, 通过主成分分析, 提取观测数据矩阵的主成分, 去除观测矩阵中存有的线性相关, 通过近似最小一乘回归得到模型参数的唯一解. 采用近似最小一乘准则, 有效地克服了大幅值尖峰序列噪声对辨识结果的影响, 提高了辨识算法的稳健性, 由于目标函数可导, 减小了算法计算量. 仿真实验表明, 当测量数据受到满足 $S\alpha S$ 分布尖峰噪声干扰时, 利用本文算法可以得到稳健的辨识结果.

参考文献 (References):

- [1] GUSTAVSSON I, LJUNG L, SÖDERSTRÖM T. Identification of processes in the closed-loop-identifiability and accuracy aspects [J]. *Automatica*, 1977, 13(4): 59 – 75.
- [2] LJUNG L. *System Identification: Theory for The Users* [M]. Englewood Cliffs, NJ: Prentice-Hall Inc, 1987: 34 – 67.
- [3] SÖDERSTRÖM T, STOICA P. *System Identification* [M]. Hemel Hempstead, UK: Prentice-Hall, 1989: 13 – 49.
- [4] GEVERS M, BOMBOIS X, HILDEBRAND R, et al. Optimal experiment design for open and closed-loop system identification [J].

- Communications in information and systems*, 2011, 11(3): 197 – 224.
- [5] CHIUSO A, PICCI G. Consistency analysis of some closed-loop subspace identification methods [J]. *Automatica*, 2005, 41(3): 377 – 391.
- [6] PIPEIKIS R. Closed-loop robust identification using the indirect approach [J]. *Information*, 2000, 11(3): 297 – 310.
- [7] MEDIANU S, LUPU C, POPESCU D, et al. Optimal closed-loop identification solution using direct and indirect methods for a flexible transmission system[C] // *The 2nd International Conference on Systems and Computer Science*. New York: IEEE, 2013: 74 – 79.
- [8] CALANCA A, L. CAPISANI M, FERRARA A, et al. MIMO closed-loop identification of an industrial robot [J]. *IEEE Transactions on Control Systems Technology*, 2011, 19(5): 1214 – 1224.
- [9] MO Jianlin, WANG Wei, XU Xiaoming, et al. Closed-loop problem in system identification [J]. *Control Theory & Applications*, 2002, 19(1): 9 – 14.
(莫建林, 王伟, 许晓鸣, 等. 系统辨识中的闭环问题 [J]. 控制理论与应用, 2002, 19(1): 9 – 14.)
- [10] FANG Chongzhi, XIAO Deyun. *Process Identification* [M]. Beijing: Tsinghua University Press, 1988: 361 – 396.
(方崇智, 萧德云. 过程辨识 [M]. 北京: 清华大学出版社, 1988: 361 – 396.)
- [11] DING F, CHEN T. Identification of hammerstein nonlinear ARMAX systems [J]. *Automatica*, 2005, 41(9): 1479 – 1489.
- [12] LÜ You, LIU Jizhen, YANG Tingting, et al. NO_x emission characteristic modeling based on feature extraction using PLS and LS-SVM [J]. *Chinese Journal of Scientific Instrument*, 2013, 34(11): 2418 – 2424.
(吕游, 刘吉臻, 杨婷婷, 等. 基于PLS特征提取和LS-SVM结合的NO_x排放特性建模 [J]. 仪器仪表学报, 2013, 34(11): 2418 – 2424.)
- [13] SUN Maojie, YANG Huizhong. Local weighted mixed kernel partial least squares algorithm and its applications to soft-sensing [J]. *Information and Control*, 2015, 44(4): 481 – 486.
(孙茂伟, 杨慧中. 局部加权混合核偏最小二乘算法及其在软测量中的应用 [J]. 信息与控制, 2015, 44(4): 481 – 486.)
- [14] WANG Wei, CHAI Tianyou, ZHAO Lijie. Dynamic partial least squares modeling with recurrent neural networks of stable learning [J]. *Control Theory & Applications*, 2012, 29(3): 337 – 341.
(王巍, 柴天佑, 赵立杰. 带有稳定学习的递归神经网络动态偏最小二乘建模 [J]. 控制理论与应用, 2012, 29(3): 337 – 341.)
- [15] QIN S J. Statistical process monitoring: basics and beyond [J]. *Journal of Chemometrics*, 2003, 17(8/9): 480 – 502.
- [16] YIN Li, LIU Qiang. Study on the identification of the milling force parameter model based on partial least square regression and application [J]. *Mechanical Science and Technology*, 2005, 24(3): 269 – 272.
(尹力, 刘强. 基于偏最小二乘回归(PLSR)方法的铣削力模型系数辨识研究 [J]. 机械科技与技术, 2005, 24(3): 269 – 272.)
- [17] ANTOINE N, ALI Ç. PLS, balanced, and canonical variate realization techniques for identifying VARMA models in state space [J]. *Chemometrics and Intelligent Laboratory Systems*, 1997, 38(2): 209 – 221.
- [18] LAURÍ D, SALCEDO J, GARCÍA-NIETO S, et al. A PLS approach to identifying predictive ARX models [C] // *The 18th IEEE International Conference on Control Applications Part of 2009 IEEE multi-conference on Systems and Control*. New York: IEEE, 2009: 1460 – 1465.
- [19] XU B C, LIU X L. Research on identification algorithm based on the approximate least absolute deviation criteria [J]. *International Journal of Automation and Computing*, 2012, 9(5): 501 – 505.
- [20] XU Baochang, ZHANG Yingdan. Multi-innovation identification algorithm based on approximate least absolute deviation [J]. *Control Engineering of China*, 2015, 22(1): 60 – 65.
(徐宝昌, 张瀛丹. 基于近似最小一乘准则的多新息辨识算法 [J]. 控制工程, 2015, 22(1): 60 – 65.)
- [21] XIANG Wei, CHEN Zonghai. New identification method of nonlinear systems based on hammerstein models [J]. *Control Theory & Applications*, 2007, 24(1): 143 – 147.
(向微, 陈宗海. 基于Hammerstein模型描述的非线性系统辨识新方法 [J]. 控制理论与应用, 2007, 24(1): 143 – 147.)
- [22] DONG Jian, XIE Kaigui. Research of the non-linear regress models based on the least absolute criteria [J]. *Journal of Chongqing Normal University (Natural Science Edition)*, 2001, 18(4): 71 – 74.
(董建, 谢开贵. 基于最小一乘准则的非线性回归模型研究 [J]. 重庆师范学院学报(自然科学版), 2001, 18(4): 71 – 74.)
- [23] DONG Qi. Relaxation algorithm based on “least absolute deviation” criterion [J]. *Predictive Method Research*, 1990, 9(2): 16 – 18.
(董祺. “残差绝对值和最小”准则的松弛算法 [J]. 预测方法研究, 1990, 9(2): 16 – 18.)
- [24] WANG Guizeng, YE Hao. *Principal Component Analysis and Partial Least Squares Method* [M]. Beijing: Tsinghua University Press, 2012: 41 – 59.
(王桂增, 叶昊. 主元分析与偏最小二乘法 [M]. 北京: 清华大学出版社, 2012: 41 – 59.)
- [25] QIN S J. Recursive PLS algorithms for adaptive data modeling [J]. *Computers & Chemical Engineering*, 1998, 22(4/5): 503 – 514.
- [26] COATES M J, KURUOGLU E E. Time-frequency based detection in impulsive noise environments using alpha-stable noise model [J]. *Digital Signal Processing*, 2002, 82(3): 1917 – 1925.
- [27] GUO Ying. *The study on novel time delay estimation methods based on stable distribution* [D]. Dalian: Dalian University of Technology, 2009.
(郭莹. 稳定分布环境下的时延估计新方法研究 [D]. 大连: 大连理工大学, 2009.)
- [28] CHEN Xiru, WU Yuehua. A problem about the relationship between the minimum L_1 -norm estimate and linear programming [J]. *Journal of Nanjing Institute of Technology*, 1988, 18(5): 16 – 21.
(陈希孺, 吴月华. 最小 L_1 模估计与线性规划的关系的一个问题 [J]. 南京工业学院学报, 1988, 18(5): 16 – 21.)

作者简介:

徐宝昌 (1974–), 男, 副教授, 硕士生导师, 目前研究方向为复杂系统建模与先进控制、多传感器信息融合与软测量技术、钻井工程自动化, E-mail: xbcyl@163.com;

林忠华 (1989–), 男, 硕士研究生, 目前研究方向为系统辨识, E-mail: linzhonghua96@163.com;

肖玉月 (1992–), 女, 硕士研究生, 目前研究方向为系统辨识, E-mail: xiaoyuyue0806@126.com.