一种 Hammerstein-Wiener 系统的递归辨识算法

于 丰[1] 毛志忠[1] 贾明兴[1] 袁 平[1] 杨飞生[2]

摘 要 针对含有过程噪声的 Hammerstein-Wiener 系统, 本文提出一种递归辨识算法用于系统的在线辨识. 首先使用多项式函数对系统非线性部分进行严格参数化, 在此基础上以参数误差平方和的期望值最小为目标函数, 推导出参数估计的递归更新公式, 避免了过程噪声对辨识结果的影响. 通过对算法进行深入分析, 得到参数一致收敛的条件, 并给出算法中重要系数的设定方法, 使参数收敛域得到扩大. 与传统两阶段法的数值仿真比较验证了该方法的优越性.

关键词 Hammerstein-Wiener 系统, 参数化, 递归辨识, 一致收敛性

引用格式 于丰, 毛志忠, 贾明兴, 袁平, 杨飞生. 一种 Hammerstein-Wiener 系统的递归辨识算法. 自动化学报, 2014, **40**(2): 327−335

DOI 10.3724/SP.J.1004.2014.00327

Recursive Identification Method for a Class of Hammerstein-Wiener Systems

YU Feng[1] MAO Zhi-Zhong[1] JIA Ming-Xing[1] YUAN Ping[1] YANG Fei-Sheng[2]

Abstract A recursive algorithm is presented to identify the Hammerstein-Wiener system with process noise. Based on parameterizing the nonlinear parts of system using polynomial functions strictly, the optimal recursive update formulas are derived in a sense that the expectation of the sum of square of parameter errors is minimized, which avoids the interference of noise. Uniform convergence conditions together with a coefficient setting method, which expands the convergence domain, are given by means of analyzing the algorithm deeply. Simulation results validate the advantage of this algorithm over the two-stage algorithm.

Key words Hammerstein-Wiener system, parameterizations, recursive identification, uniform convergence

Citation Yu Feng, Mao Zhi-Zhong, Jia Ming-Xing, Yuan Ping, Yang Fei-Sheng. Recursive identification method for a class of Hammerstein-Wiener systems. *Acta Automatica Sinica*, 2014, **40**(2): 327−335

许多非线性系统可以表示为线性动态模块与非线性静态模块相串联的形式, 这类系统通称为块状系统. 根据各模块串联顺序的不同, 可以分为四种不同形式的模型. 其中, 由非线性静态模块串联线性动态模块组成的系统称为 Hammerstein 系统 (H 系统); 由线性动态模块串联非线性静态模块组成的系统则称为 Wiener 系统 (W 系统). 这两种系统相组合可以得到 Hammerstein-Wiener 系统 (H-W 系统) 以及 Wiener-Hammerstein 系统 (W-H 系统)[1].

对于块状系统的辨识算法研究, 已经得到许多成果, 其中对于 H 系统及 W-H 系统的研究较为成熟[2−7]; 对于 W 系统辨识方法的研究则相对不足[8−10], 这主要由于缺乏有效表示输出非线性的方法[11]; 而对于模型结构进一步复杂的 H-W 系统, 现有的研究成果更为稀少, 其中最有代表性的是文献 [12] 提出的两阶段辨识法, 该方法将输入非线性、输出非线性逆表示成多项式函数, 第一阶段通过最小二乘法 (Least square, LS) 得到原系统参数乘积项; 第二阶段使用奇异值分解法 (Singular value decomposition, SVD) 分离线性部分与非线性部分参数. 根据其作者所述, 这类方法的缺陷在于文献中式 (1) 是对 H-W 系统的近似参数化表示, 它忽略了输出非线性逆多项式函数在 t 时刻的高次项; 同时当噪声作用于系统时, 前一阶段 LS 的辨识误差会导致第二阶段 SVD 的结果偏离真值. 在线辨识过程中, 这种估计误差表现得更为明显. 文献 [13−15] 在两阶段辨识法的基础上提出了改进方法, 但仍没有从本质上解决该方法的不足. 文献 [11] 在文献 [12] 的基础上提出了用于在线辨识的盲方法, 其特殊的采样方式限制了算法的应用. 文献 [16] 提出一种松弛迭代法; 文献 [17] 分别使用正交基函数以及样条函数对系统线性与非线性部分进行参数化, 进而对系统进行辨识; 文献 [18] 使用一种特殊神经网络对 H-W 系统进行辨识. 以上三种方法计算量较大, 只

收稿日期 2012-07-10 录用日期 2013-01-06
Manuscript received July 10, 2012; accepted January 6, 2013
国家自然科学基金 (61074098, 61203103, 61333006), 中央高校基本科研业务费 (N1103040 06) 资助
Supported by National Natural Science Foundation of China (61074098, 61203103, 61333006) and Fundamental Research Funds for the Central Universities (N110304006)
本文责任编委 方海涛
Recommended by Associate Editor FANG Hai-Tao
1. 东北大学信息科学与工程学院 沈阳 110819 2. 西北工业大学自动化学院 西安 710072
1. School of Information Science and Engineering, Northeastern University, Shenyang 110819 2. School of Automation, Northwestern Polytechnical University, Xi′an 710072

适用于离线辨识. 可见对于 H-W 系统的辨识, 特别是在线辨识算法的研究还远远不够.

鉴于现有工作的缺乏与不足, 本文针对存在过程噪声的 H-W 系统, 根据扩展 Kalman 滤波 (Extended Kalman filter, EKF) 算法的推导思路, 提出一种用于在线辨识 H-W 系统的递归算法 (MEKF). 不同于文献 [12] 的近似处理, 该方法对系统进行了严格的参数化. 在算法推导过程中, 以参数误差平方和的期望值最小为目标函数, 避免了过程噪声对参数估计的影响. 通过对算法的分析, 得到参数一致收敛的条件. 在此基础上给出算法中关键系数的设定方法, 使参数估计的收敛域得到扩大.

本文的结构安排如下: 第 1 节对 H-W 系统进行参数化表示; 第 2 节在对真实参数以及参数归一化问题进行说明的基础上, 推导了参数的递归更新公式; 第 3 节证明了算法的一致收敛性, 提出相应条件并给出可以扩大算法收敛域的算法参数设定方法; 第 4 节通过两个仿真例子表明算法的有效性; 第 5 节总结全文, 得出结论.

1 H-W 模型参数化

H-W 系统结构如图 1 所示.

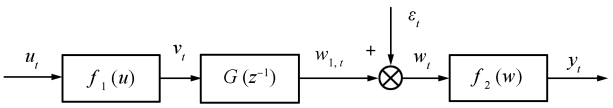


图 1 H-W 系统

Fig. 1 Structure diagram of Hammerstein-Wiener system

图 1 中, u_t 和 y_t 分别为系统输入与输出; v_t, $w_{1,t}$, w_t 为中间过程变量; ε_t 为过程噪声, 作用于系统中间过程变量 $w_{1,t}$. H-W 系统各模块模型可以由式 (1) ∼ (3) 表示:

$$v_t = f_1(u_t) \tag{1}$$

$$w_t = G(z^{-1})v_t + \varepsilon_t \tag{2}$$

$$y_t = f_2(w_t) \tag{3}$$

其中, $f_1(u_t)$ 表示系统输入非线性; $G(z^{-1})$ 表示线性动态部分离散传递函数, z^{-1} 为单位延迟符号; $f_2(w_t)$ 表示系统输出非线性.

首先需要对 H-W 系统作如下假设:

假设 1. 输入输出非线性 $f_1(u_t)$, $f_2(w_t)$ 是连续的, $f_2(w_t)$ 是可逆的[11−12, 15−16];

假设 2. 线性部分的阶数是已知的, 即辨识不包括对于模型线性部分的结构估计[11];

假设 3. 过程噪声 ε_t 为均值为 0、方差为 δ^2 的白噪声, 噪信比不能过大[11−13, 16, 19];

假设 4. 输出量 y_t, 参数误差 $\tilde{\boldsymbol{\theta}}_t$ 都与干扰相关, 考虑到当噪声较小时, 这种相关性很小, 因此在算法推导过程中忽略这种相关性[19].

在推导算法前, 要对 H-W 模型进行参数化, 即使用结构已知的数学模型对原系统进行近似表示. 对于 H-W 系统中输入输出非线性特性, 可以使用已知的非线性基函数的线性组合来描述. 本文与文献 [12−13] 相同, 使用多项式函数对输入非线性进行近似描述:

$$v_t = f_1(u_t) = \sum_{i=1}^{p} c_i u_t^i \tag{4}$$

其中, c_i 为未知系数, p 为多项式函数阶数.

算法需要输出非线性逆的参数化表示, 因此使用多项式函数对输出非线性逆函数进行近似描述:

$$w_t = f_2^{-1}(y_t) = \sum_{i=1}^{q} d_i y_t^i \tag{5}$$

其中, d_i 为未知系数, q 为多项式函数阶数.

对于线性动态部分, 可以使用与原系统结构相同的离散传递函数模型进行表示:

$$G(z^{-1}) = \frac{b_1 z^{-1} + b_2 z^{-2} + \cdots + b_m z^{-m}}{a_1 + a_2 z^{-1} + \cdots + a_r z^{-r+1}} \tag{6}$$

其中, a_i, b_i 为模型参数; m, r 为模型阶数. 令:

$$\boldsymbol{a} = [a_1, a_2, \cdots, a_r]^{\mathrm{T}},\ \ \boldsymbol{b} = [b_1, b_2, \cdots, b_m]^{\mathrm{T}}$$

$$\boldsymbol{c} = [c_1, c_2, \cdots, c_p]^{\mathrm{T}},\ \ \boldsymbol{d} = [d_1, d_2, \cdots, d_q]^{\mathrm{T}}$$

$$\boldsymbol{y}_t = [y_t, y_t^2, \cdots, y_t^q]^{\mathrm{T}},\ \ \boldsymbol{u}_t = [u_t, u_t^2, \cdots, u_t^p]^{\mathrm{T}}$$

$$\boldsymbol{\varepsilon}_t = [\varepsilon_t, \varepsilon_{t-1}, \cdots, \varepsilon_{t-r+1}]^{\mathrm{T}},\ \boldsymbol{\theta} = [\boldsymbol{a}^{\mathrm{T}}, \boldsymbol{b}^{\mathrm{T}}, \boldsymbol{d}^{\mathrm{T}}, \boldsymbol{c}^{\mathrm{T}}]^{\mathrm{T}}$$

$$Y_{rq,t} = \begin{bmatrix} y_t & y_t^2 & \cdots & y_t^q \\ y_{t-1} & y_{t-1}^2 & \cdots & y_{t-1}^q \\ \vdots & \vdots & \ddots & \vdots \\ y_{t-r+1} & y_{t-r+1}^2 & \cdots & y_{t-r+1}^q \end{bmatrix}$$

$$U_{mp,t} = \begin{bmatrix} u_{t-1} & u_{t-1}^2 & \cdots & u_{t-1}^p \\ u_{t-2} & u_{t-2}^2 & \cdots & y_{t-2}^p \\ \vdots & \vdots & \ddots & \vdots \\ u_{t-m} & u_{t-m}^2 & \cdots & u_{t-m}^p \end{bmatrix}$$

并定义当前参数估计值为

$$\hat{\boldsymbol{\theta}}_t = \left[\hat{\boldsymbol{a}}_t^{\mathrm{T}}, \hat{\boldsymbol{b}}_t^{\mathrm{T}}, \hat{\boldsymbol{d}}_t^{\mathrm{T}}, \hat{\boldsymbol{c}}_t^{\mathrm{T}}\right]^{\mathrm{T}}$$

$$\hat{\boldsymbol{a}}_t = [\hat{a}_{1,t}, \hat{a}_{2,t}, \cdots, \hat{a}_{r,t}]^{\mathrm{T}}$$

$$\hat{\boldsymbol{b}}_t = \left[\hat{b}_{1,t}, \hat{b}_{2,t}, \cdots, \hat{b}_{m,t}\right]^{\mathrm{T}}$$

$$\hat{\boldsymbol{c}}_t = [\hat{c}_{1,t}, \hat{c}_{2,t}, \cdots, \hat{c}_{p,t}]^{\mathrm{T}}$$

$$\hat{\boldsymbol{d}}_t = \left[\hat{d}_{1,t}, \hat{d}_{2,t}, \cdots, \hat{d}_{q,t}\right]^{\mathrm{T}}$$

其中, $\hat{a}_{i,t}$, $\hat{b}_{i,t}$, $\hat{c}_{i,t}$, $\hat{d}_{i,t}$ 分别为 a_i, b_i, c_i, d_i 在 t 时刻的估计值. 经过参数化后的 H-W 系统可以表示为如下形式:

$$\boldsymbol{a}^{\mathrm{T}} Y_{rq,t} \boldsymbol{d} = \boldsymbol{b}^{\mathrm{T}} U_{mp,t} \boldsymbol{c} + \boldsymbol{a}^{\mathrm{T}} \boldsymbol{\varepsilon}_t \tag{7}$$

2　参数估计递归算法

在对系统进行参数化表示的基础上, 根据 EKF 算法的思路, 推导估计参数的递归更新公式, 公式推导的主要步骤为: 首先推导参数误差 $\tilde{\boldsymbol{\theta}}_t$ 的近似表达式, 进而推导出参数误差协方差期望矩阵 $\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t \tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]$, 令该矩阵的迹, 即参数误差平方和的期望值 $\sum \mathrm{E}\left[\tilde{\boldsymbol{\theta}}_i^2\right]$ 最小, 得到参数递归更新率. 基于此, 首先对系统的真实参数进行说明.

2.1　原系统真实参数

由于算法通过令参数误差平方和的期望值最小得到辨识参数更新率, 因此原 H-W 系统中应该存在与系统参数化模型对应的真实参数, 即存在参数向量 $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$, $\boldsymbol{d}$, 使式 (7) 体现原系统特性.

对于系统线性部分, 参数化模型所取结构与原系统相同, 因此存在真实参数向量 $\boldsymbol{a}$, $\boldsymbol{b}$; 而输入非线性以及输出非线性逆特性结构未知, 在参数化过程中使用多项式函数近似表示, 原系统中不存在真实参数向量 $\boldsymbol{c}$, $\boldsymbol{d}$. 但是对于给定的多项式阶数 p, q, 在非线性特性输入输出范围内, 一定存在一组参数向量使多项式函数与原非线性特性拟合误差最小, 此时认为该组参数为原系统非线性特性的真实参数向量.

在块状结构参数辨识方法研究中, 参数一致性问题是不可避免的, 即考虑到干扰的不可测性, 任何辨识算法都无法辨认式 (7) 中参数对 ($\boldsymbol{a}$, $\boldsymbol{d}$) 与 ($n\boldsymbol{a}$, $n^{-1}\boldsymbol{d}$), 或参数对 ($\boldsymbol{b}$, $\boldsymbol{c}$) 与 ($n\boldsymbol{b}$, $n^{-1}\boldsymbol{c}$) 的区别, 这导致了真实参数并不唯一. 解决这个问题的一个简单有效的方法是令 $\hat{a}_{1,t}$, $\hat{c}_{1,t}$, $\hat{d}_{1,t}$ 为恒值 1, 不参与参数递归更新过程. 此时, 可以认为真实参数向量为 $\boldsymbol{a}^{\mathrm{T}}/a_1$, $c_1\boldsymbol{b}^{\mathrm{T}}/d_1 a_1$, $\boldsymbol{c}^{\mathrm{T}}/c_1$, $\boldsymbol{d}^{\mathrm{T}}/d_1$.

2.2　算法推导

基于以上对真实参数的说明, 以及前文提出的公式推导思路, 推导参数更新公式.

定义参数估计误差:

$$\tilde{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta} \tag{8}$$

为获得良好的参数辨识及抗干扰能力, 希望参数估计值最大程度靠近真值, 可以令目标函数:

$$v_t = \mathrm{tr}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t \tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]\right) \tag{9}$$

最小.

定义当前参数估计值以及模型预测误差:

$$\hat{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_{t-1} + \boldsymbol{k}_t e_t \tag{10}$$

$$e_t = \hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}} Y_{rq,t} \hat{\boldsymbol{d}}_{t-1} - \hat{\boldsymbol{b}}_{t-1}^{\mathrm{T}} U_{mp,t} \hat{\boldsymbol{c}}_{t-1} \tag{11}$$

其中, $\boldsymbol{k}_t$ 为参数更新向量.

为了获得参数误差协方差期望矩阵 $\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t \tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]$, 首先推导参数误差 $\tilde{\boldsymbol{\theta}}_t$ 的表达式. 结合式 (8) 和 (10), 得:

$$\tilde{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_t - \boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}_{t-1} + \boldsymbol{k}_t e_t \tag{12}$$

根据式 (7)、(8) 以及 (11), 可得:

$$\begin{aligned} e_t = {} & \hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}} Y_{rq,t} \hat{\boldsymbol{d}}_{t-1} - \hat{\boldsymbol{b}}_{t-1}^{\mathrm{T}} U_{mp,t} \hat{\boldsymbol{c}}_{t-1} - \boldsymbol{a}^{\mathrm{T}} Y_{rq,t} \boldsymbol{d} + \\ & \boldsymbol{b}^{\mathrm{T}} U_{mp,t} \boldsymbol{c} + \boldsymbol{a}^{\mathrm{T}} \boldsymbol{\varepsilon}_t = \\ & \hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}} Y_{rq,t} \tilde{\boldsymbol{d}}_{t-1} + \tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}} Y_{rq,t} \hat{\boldsymbol{d}}_{t-1} - \\ & \tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}} Y_{rq,t} \tilde{\boldsymbol{d}}_{t-1} - \tilde{\boldsymbol{b}}_{t-1}^{\mathrm{T}} U_{mp,t} \hat{\boldsymbol{c}}_{t-1} - \\ & \hat{\boldsymbol{b}}_{t-1}^{\mathrm{T}} U_{mp,t} \tilde{\boldsymbol{c}}_{t-1} + \tilde{\boldsymbol{b}}_{t-1}^{\mathrm{T}} U_{mp,t} \tilde{\boldsymbol{c}}_{t-1} + \hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}} \boldsymbol{\varepsilon}_t - \\ & \tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}} \boldsymbol{\varepsilon}_t \end{aligned} \tag{13}$$

当参数估计值接近真值时, $\tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}} Y_{rq,t} \tilde{\boldsymbol{d}}_{t-1}$, $\tilde{\boldsymbol{b}}_{t-1}^{\mathrm{T}} U_{mp,t} \tilde{\boldsymbol{c}}_{t-1}$, $\tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}} \boldsymbol{\varepsilon}_t$ 项相对于模型预测误差 e_t 较小, 可以忽略, 得到 e_t 的近似表达式:

$$\begin{aligned} e_t \approx {} & \hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}} Y_{rq,t} \tilde{\boldsymbol{d}}_{t-1} + \tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}} Y_{rq,t} \hat{\boldsymbol{d}}_{t-1} + \hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}} \boldsymbol{\varepsilon}_t - \\ & \tilde{\boldsymbol{b}}_{t-1}^{\mathrm{T}} U_{mp,t} \hat{\boldsymbol{c}}_{t-1} - \hat{\boldsymbol{b}}_{t-1}^{\mathrm{T}} U_{mp,t} \tilde{\boldsymbol{c}}_{t-1} = \\ & \hat{\boldsymbol{\theta}}_{l,t-1}^{\mathrm{T}} M_t \tilde{\boldsymbol{\theta}}_{n,t-1} + \hat{\boldsymbol{\theta}}_{n,t-1}^{\mathrm{T}} M_t^{\mathrm{T}} \tilde{\boldsymbol{\theta}}_{l,t-1} + \hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}} \boldsymbol{\varepsilon}_t \end{aligned} \tag{14}$$

其中, $\boldsymbol{\theta}_l = \left(\boldsymbol{a}^{\mathrm{T}}, \boldsymbol{b}^{\mathrm{T}}\right)^{\mathrm{T}}$, $\boldsymbol{\theta}_n = \left(\boldsymbol{d}^{\mathrm{T}}, \boldsymbol{c}^{\mathrm{T}}\right)^{\mathrm{T}}$, $M_t = \begin{bmatrix} Y_{rq,t} & 0 \\ 0 & -U_{mp,t} \end{bmatrix}$.

由式 (12) 和 (14), 得到参数估计误差 $\tilde{\boldsymbol{\theta}}_t$ 的表达式:

$$\begin{aligned} \tilde{\boldsymbol{\theta}}_t = {} & \tilde{\boldsymbol{\theta}}_{t-1} + \\ & \boldsymbol{k}_t \left\{ \left[\hat{\boldsymbol{\theta}}_{n,t-1}^{\mathrm{T}} M_t^{\mathrm{T}}, \hat{\boldsymbol{\theta}}_{l,t-1}^{\mathrm{T}} M_t\right] \tilde{\boldsymbol{\theta}}_{t-1} + \hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}} \boldsymbol{\varepsilon}_t \right\} \end{aligned} \tag{15}$$

由式 (15), 在当前输入输出数据已知条件下, 可以得到参数误差协方差期望矩阵 $\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t \tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]$ 的表达式:

$$\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]=\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]+\boldsymbol{k}_t\boldsymbol{\beta}_t^{\mathrm{T}}+\boldsymbol{\beta}_t\boldsymbol{k}_t^{\mathrm{T}}+\boldsymbol{k}_t\alpha_t\boldsymbol{k}_t^{\mathrm{T}} \tag{16}$$

其中,

$$\alpha_t=\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}+R_t^2 \tag{17}$$

$$\boldsymbol{\beta}_t=\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1} \tag{18}$$

$$\boldsymbol{f}_{t-1}^{\mathrm{T}}=\left[\hat{\boldsymbol{\theta}}_{n,t-1}^{\mathrm{T}}M_t^{\mathrm{T}},\hat{\boldsymbol{\theta}}_{l,t-1}^{\mathrm{T}}M_t\right] \tag{19}$$

$$R_t^2=\mathrm{E}\left[\hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t^{\mathrm{T}}\hat{\boldsymbol{a}}_{t-1}\right] \tag{20}$$

定理 1. 实现目标函数 (9) 的近似最优参数更新向量如下式所示:

$$\boldsymbol{k}_t=-\alpha_t^{-1}\boldsymbol{\beta}_t \tag{21}$$

证明. 注意到式 (9) 中 v_t 为一实数, 对 v_t 求 $\boldsymbol{k}_t$ 的偏导数:

$$\frac{\partial\left[\mathrm{tr}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]\right)\right]}{\partial\boldsymbol{k}_t}=2\boldsymbol{\beta}_t+2\alpha_t\boldsymbol{k}_t \tag{22}$$

$\Rightarrow \boldsymbol{k}_t=-\alpha_t^{-1}\boldsymbol{\beta}_t$.

对 v_t 求 $\boldsymbol{k}_t$ 的二阶偏导数:

$$\frac{\partial^2\left[\mathrm{tr}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]\right)\right]}{\partial\boldsymbol{k}_t^2}=2\alpha_t I \tag{23}$$

其中, I 为单位矩阵.

由于 $\alpha_t>0$, $2\alpha_t I$ 正定, 式 (16) 为 $\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]$ 的近似表达, 因此式 (21) 可以作为保证目标函数 (9) 达到最小的近似最优解. □

在式 (14) 和 (16) 以及定理 1 的证明过程中可知, 这种近似最优是以参数误差较小为基础的, 当参数误差较大时, 式 (21) 为次优的参数更新向量.

3 参数一致收敛性分析

一致收敛性分析是参数辨识算法中重要的组成部分, 接下来在文献 [19] 对 EKF 收敛性分析的基础上对该辨识算法进行分析, 得到一致收敛条件以及一些有意义的结论.

3.1 参数一致收敛性分析

定义符号 $P_{(3)}$: 当 $\boldsymbol{P}$ 为向量时, $\boldsymbol{P}_{(3)}$ 表示向量 $\boldsymbol{P}$ 去掉第 1, $r+m+1$, $r+m+q+1$ 个元素后剩余的子向量; 当 P 为矩阵时, $P_{(3)}$ 表示矩阵 P 去掉第 1, $r+m+1$, $r+m+q+1$ 行、列后剩余的子矩阵.

定理 2. 当满足:

$$-\sqrt{1-\frac{\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}}{R_t^2}}\le\mu_t\le\sqrt{1-\frac{\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}}{R_t^2}} \tag{24}$$

以及当 R_t^2 为有限正实数, $\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^{\mathrm{T}}\right]$ 为有限正定阵, 且存在有限正实数 η_1, η_2 和某一有限整数 $N\ge1$, 对于任意 $t\ge N$, 使:

$$\eta_1 I\le\sum_{i=t-N+1}^{t}\boldsymbol{f}_{i(3)}\boldsymbol{f}_{i(3)}^{\mathrm{T}}\le\eta_2 I \tag{25}$$

满足时, 使用该算法调整模型参数, 可以使

$$\lim_{t\to\infty}\tilde{\boldsymbol{\theta}}_t=0 \tag{26}$$

其中

$$\mu_t=(-\tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}}Y_{rq,t}\tilde{\boldsymbol{d}}_{t-1}+\tilde{\boldsymbol{b}}_{t-1}^{\mathrm{T}}U_{mp,t}\tilde{\boldsymbol{c}}_{t-1}+\hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}}\boldsymbol{\varepsilon}_t-\tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}}\boldsymbol{\varepsilon}_t)e_t^{-1} \tag{27}$$

证明. 将式 (21) 代入式 (16), 得:

$$\begin{aligned}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]=&\ \mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]-\\&\left[\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}+R_t^2\right]^{-1}\times\\&\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\end{aligned} \tag{28}$$

由式 (28) 得到:

$$\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]=\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]^{-1}+\boldsymbol{f}_{t-1}R_t^2\boldsymbol{f}_{t-1}^{\mathrm{T}} \tag{29}$$

$$\begin{aligned}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}=&\ R_t^2\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}\times\\&\left[\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}+R_t^2\right]^{-1}\end{aligned} \tag{30}$$

定义 Lyapunov 函数:

$$l_t=\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]^{-1}\tilde{\boldsymbol{\theta}}_t \tag{31}$$

将式 (21) 代入式 (12), 再代入式 (31), 得:

$$\begin{aligned}l_t=&\{\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}-\left[\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}+R_t^2\right]^{-1}\times\\&\boldsymbol{f}_{t-1}^{\mathrm{T}}\ \mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]e_t\}\cdot\ \mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]^{-1}\times\\&\{\tilde{\boldsymbol{\theta}}_{t-1}-\left[\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}+R_t^2\right]^{-1}\times\\&\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}e_t\}\end{aligned} \tag{32}$$

将式 (29) 和 (30) 代入式 (32), 得:

$$l_t = l_{t-1} + R_t^{-2}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\boldsymbol{f}_{t-1}\boldsymbol{f}_{t-1}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}_{t-1} - 2R_t^{-2}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\boldsymbol{f}_{t-1}e_t + R_t^{-4}\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}e_t^2 \tag{33}$$

由式 (13), 得:

$$\boldsymbol{f}_{t-1}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}_{t-1} = e_t + \tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}}Y_{rq,t}\tilde{\boldsymbol{d}}_{t-1} - \tilde{\boldsymbol{b}}_{t-1}^{\mathrm{T}}U_{mp,t}\tilde{\boldsymbol{c}}_{t-1} - \hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}}\boldsymbol{\varepsilon}_t + \tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}}\boldsymbol{\varepsilon}_t \tag{34}$$

根据式 (14) 的推导过程, 忽略了 $\tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}}Y_{rq,t}\tilde{\boldsymbol{d}}_{t-1}$, $\tilde{\boldsymbol{b}}_{t-1}^{\mathrm{T}}U_{mp,t}\tilde{\boldsymbol{c}}_{t-1}$, $\tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}}\boldsymbol{\varepsilon}_t$ 项. 为了准确分析算法的收敛性, 引入系数 μ_t, 令:

$$\mu_t e_t = -\tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}}Y_{rq,t}\tilde{\boldsymbol{d}}_{t-1} + \tilde{\boldsymbol{b}}_{t-1}^{\mathrm{T}}U_{mp,t}\tilde{\boldsymbol{c}}_{t-1} + \hat{\boldsymbol{a}}_{t-1}^{\mathrm{T}}\boldsymbol{\varepsilon}_t - \tilde{\boldsymbol{a}}_{t-1}^{\mathrm{T}}\boldsymbol{\varepsilon}_t \tag{35}$$

将式 (35) 代入式 (34), 有:

$$e_t = \boldsymbol{f}_{t-1}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}_{t-1} + \mu_t e_t \tag{36}$$

可见系数 μ_t 表达了式 (14) 中忽略项以及未知干扰项在模型误差 e_t 中的比重关系. 将式 (36) 代入式 (33), 得:

$$l_t = l_{t-1} + \left[R_t^{-2}(1-\mu_t)^2 - 2R_t^{-2}(1-\mu_t) + R_t^{-4}\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}\right]e_t^2 \tag{37}$$

为了使 $l_t \le l_{t-1}$, 需保证:

$$R_t^{-2}(1-\mu_t)^2 - 2R_t^{-2}(1-\mu_t) + R_t^{-4}\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]\boldsymbol{f}_{t-1} \le 0 \tag{38}$$

解不等式方程, 可得条件 (24). 此时, l_t 值随时间 t 逐渐减小, 即:

$$0 \le l_t \le l_{t-1} \le \cdots \le l_0 \tag{39}$$

参数估计值逐渐收敛于一定值.

接下来分析参数收敛于真值的条件. 由于设定 $\hat{a}_{1,t}$, $\hat{c}_{1,t}$, $\hat{d}_{1,t}$ 为恒值 1, 因此在下面的讨论中不再将这 3 个元素纳入考虑范围.

由条件 (25) 知, 对于整数 $k \ge 1$, 当 $t \ge kN$ 时, 有:

$$k\eta_1 I \le \sum_{i=t-kN+1}^{t}\boldsymbol{f}_{i(3)}\boldsymbol{f}_{i(3)}^{\mathrm{T}} \le k\eta_2 I \tag{40}$$

当 $k \to \infty$ 时, 亦有 $t \to \infty$, 因此根据式 (40) 可得:

$$\lambda_{\min}\left(\lim_{t\to\infty}\sum_{i=1}^{t}(\boldsymbol{f}_{i(3)}\boldsymbol{f}_{i(3)}^{\mathrm{T}})\right) = \lambda_{\min}\left(\lim_{k\to\infty}\sum_{i=0}^{k-1}\sum_{t=iN+1}^{iN+N}(\boldsymbol{f}_{t(3)}\boldsymbol{f}_{t(3)}^{\mathrm{T}})\right) \ge \lim_{k\to\infty}\sum_{i=0}^{k-1}\lambda_{\min}\left(\sum_{t=iN+1}^{iN+N}(\boldsymbol{f}_{t(3)}\boldsymbol{f}_{t(3)}^{\mathrm{T}})\right) \ge \lim_{k\to\infty}k\eta_1 = \infty \tag{41}$$

$$\frac{\lambda_{\max}\left(\lim_{t\to\infty}\sum_{i=1}^{t}(\boldsymbol{f}_{i(3)}\boldsymbol{f}_{i(3)}^{\mathrm{T}})\right)}{\lambda_{\min}\left(\lim_{t\to\infty}\sum_{i=1}^{t}(\boldsymbol{f}_{i(3)}\boldsymbol{f}_{i(3)}^{\mathrm{T}})\right)} = \frac{\lambda_{\max}\left(\lim_{k\to\infty}\sum_{i=0}^{k-1}\sum_{t=iN+1}^{iN+N}(\boldsymbol{f}_{t(3)}\boldsymbol{f}_{t(3)}^{\mathrm{T}})\right)}{\lambda_{\min}\left(\lim_{k\to\infty}\sum_{i=0}^{k-1}\sum_{t=iN+1}^{iN+N}(\boldsymbol{f}_{t(3)}\boldsymbol{f}_{t(3)}^{\mathrm{T}})\right)} \le \frac{\lim_{k\to\infty}\sum_{i=0}^{k-1}\lambda_{\max}\left(\sum_{t=iN+1}^{iN+N}(\boldsymbol{f}_{t(3)}\boldsymbol{f}_{t(3)}^{\mathrm{T}})\right)}{\lim_{k\to\infty}\sum_{i=0}^{k-1}\lambda_{\min}\left(\sum_{t=iN+1}^{iN+N}(\boldsymbol{f}_{t(3)}\boldsymbol{f}_{t(3)}^{\mathrm{T}})\right)} \le \lim_{k\to\infty}\frac{k\eta_2}{k\eta_1} = \frac{\eta_2}{\eta_1} < \infty \tag{42}$$

其中, $\lambda(\cdot)$ 表示矩阵的特征值.

根据式 (29), 可得:

$$\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1} = \mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]_{(3)}^{-1} + R_t^2\boldsymbol{f}_{t-1(3)}\boldsymbol{f}_{t-1(3)}^{\mathrm{T}} = \mathrm{E}\left[\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^{\mathrm{T}}\right]_{(3)}^{-1} + \sum_{i=0}^{t-1}R_{i+1}^2\boldsymbol{f}_{i(3)}\boldsymbol{f}_{i(3)}^{\mathrm{T}} \tag{43}$$

考虑到 $\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}$ 为正定对称阵以及 R_t^2 为有限正实数、$\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^{\mathrm{T}}\right]$ 为有限正定阵, 因此有:

$$\lambda_{\min}\lim_{t\to\infty}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}\right) \ge \lambda_{\min}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^{\mathrm{T}}\right]_{(3)}^{-1}\right) + \lambda_{\min}\left(\lim_{t\to\infty}\sum_{i=0}^{t-1}R_{i+1}^2\boldsymbol{f}_{i(3)}\boldsymbol{f}_{i(3)}^{\mathrm{T}}\right) = \infty \tag{44}$$

$$1 < \lim_{t\to\infty}\frac{\lambda_{\max}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}\right)}{\lambda_{\min}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}\right)} \le$$

$$\frac{\lambda_{\max}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^{\mathrm{T}}\right]_{(3)}^{-1}\right)+\lambda_{\max}\left(\lim\limits_{t\to\infty}\sum\limits_{i=0}^{t-1}R_{i+1}^2\boldsymbol{f}_{i(3)}\boldsymbol{f}_{i(3)}^{\mathrm{T}}\right)}{\lambda_{\min}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^{\mathrm{T}}\right]_{(3)}^{-1}\right)+\lambda_{\min}\left(\lim\limits_{t\to\infty}\sum\limits_{i=0}^{t-1}R_{i+1}^2\boldsymbol{f}_{i(3)}\boldsymbol{f}_{i(3)}^{\mathrm{T}}\right)}<\infty \tag{45}$$

定义实数 ω:

$$\omega=\lim_{t\to\infty}\frac{\lambda_{\max}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}\right)}{\lambda_{\min}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}\right)} \tag{46}$$

其中, $1<\omega<\infty$, 以及正常数 $l_{(3)}$

$$\lim_{t\to\infty}l_{t(3)}=l_{(3)} \tag{47}$$

其中,

$$l_{t(3)}=\tilde{\boldsymbol{\theta}}_{t(3)}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}\tilde{\boldsymbol{\theta}}_{t(3)} \tag{48}$$

定义正数 r_t:

$$r_t=\mathrm{tr}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}\right) \tag{49}$$

根据式 (44), 知:

$$\lim_{t\to\infty}r_t=\infty \tag{50}$$

结合式 (48), 可得:

$$\lim_{t\to\infty}\frac{l_{t(3)}}{r_t}=0 \tag{51}$$

因此根据式 (46), 有:

$$\begin{aligned}0=\lim_{t\to\infty}\frac{l_{t(3)}}{r_t}\geq\\ \lim_{t\to\infty}\frac{\lambda_{\min}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}\right)\tilde{\boldsymbol{\theta}}_{t(3)}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}_{t(3)}}{r_t}\geq\\ \lim_{t\to\infty}\frac{\lambda_{\min}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}\right)\tilde{\boldsymbol{\theta}}_{t(3)}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}_{t(3)}}{(r+m+p+q-3)\lambda_{\max}\left(\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]_{(3)}^{-1}\right)}\geq\\ \lim_{t\to\infty}\frac{\tilde{\boldsymbol{\theta}}_{t(3)}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}_{t(3)}}{(r+m+p+q-3)\omega}\end{aligned} \tag{52}$$

可得:

$$\lim_{t\to\infty}\tilde{\boldsymbol{\theta}}_{t(3)}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}_{t(3)}=0 \tag{53}$$

即:

$$\lim_{t\to\infty}\tilde{\boldsymbol{\theta}}_t=0$$

□

由于条件 (25) 对 $\boldsymbol{f}_{i(3)}\boldsymbol{f}_{i(3)}^{\mathrm{T}}$ 进行了限制, 而 $\boldsymbol{f}_{i(3)}$ 中包含了输入输出数据, 因此条件 (25) 可以作为算法充分激励的条件.

在保证一致收敛的基础上, 如何扩大参数的收敛域是值得研究的问题, 条件 (24) 为解决这一问题提供了依据. 由式 (24) 可知, μ_t 的取值范围决定算法收敛域的大小, 而式中唯一可变量 R_t^2 值的设定方法与 μ_t 的取值范围密切相关, 可以通过改变 R_t^2 值以实现收敛域的扩大. 通过式 (20) 可知, R_t^2 值本应该与噪声方差有关, 而作为一种基于 EKF 的辨识方法, 与 EKF 相同, 在算法实际应用过程中, R_t^2 值并不一定需要按照式 (20) 进行计算[20−21], 适当改变 R_t^2 值的设定方法有可能会提高算法的辨识能力[22−23]. 这里借鉴文献 [19] 的研究成果, 令:

$$R_t^2=n\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}+c \tag{54}$$

其中, $n>1$, c 为大于 0 的常数. 当 n 取一较大值时, μ_t 的范围随之扩大, 进而扩大了收敛域. 取 c 为大于 0 的常数是为防止算法出现震荡.

需要指出的是, 参数实现一致收敛并不一定能够获得良好的模型预测输出. 这是由于参数辨识算法不能克服输入输出非线性参数化结构不合理带来的误差. 举一个简单的例子: 使用低阶多项式函数对一复杂非线性特性进行近似描述, 使用任何参数组都无法达到满意的拟合效果. 解决这一问题只能依靠改变模型结构来实现. 在实际应用过程中, 当参数收敛后, 如果模型预测误差较大, 可以考虑适当增加多项式函数阶数以提高模型精度.

3.2 算法步骤与初始化

首先给出参数更新递归公式:

$$\boldsymbol{k}_t=-\alpha_t^{-1}\boldsymbol{\beta}_t \tag{55}$$

$$\hat{\boldsymbol{\theta}}_t=\hat{\boldsymbol{\theta}}_{t-1}+\boldsymbol{k}_te_t \tag{56}$$

$$\begin{aligned}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_t\tilde{\boldsymbol{\theta}}_t^{\mathrm{T}}\right]=\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]+\\ \boldsymbol{k}_t\boldsymbol{\beta}_t^{\mathrm{T}}+\boldsymbol{\beta}_t\boldsymbol{k}_t^{\mathrm{T}}+\boldsymbol{k}_t\alpha_t\boldsymbol{k}_t^{\mathrm{T}}\end{aligned} \tag{57}$$

接着给出算法初始化方法:

1) 给出估计参数初始值 $\hat{\boldsymbol{\theta}}_0$. 根据参数归一化原则, 令 $\hat{a}_0$, $\hat{c}_0$, $\hat{d}_0$ 值为 1. 如果对原系统有一定的先验知识, $\hat{\boldsymbol{\theta}}_0$ 中其他值应尽量接近参数真值; 否则, 可以取:

$$\hat{\boldsymbol{\theta}}_0=(\underbrace{1,0,\cdots,0}_{r+m},\underbrace{1,0,\cdots,0}_{q},\underbrace{1,0,\cdots 0}_{p})^{\mathrm{T}} \tag{58}$$

2) 初始化参数误差协方差期望矩阵 $\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^{\mathrm{T}}\right]$. 由于 $\hat{a}_{1,t}$, $\hat{c}_{1,t}$, $\hat{d}_{1,t}$ 为恒值, 不参与更新, 因此令

$\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^{\mathrm{T}}\right]$ 中第 (1, 1), $(r+m+1, r+m+1)$ 与 $(r+m+q+1, r+m+q+1)$ 元素值为 0. 矩阵初始值可以按下式给出:

$$\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^{\mathrm{T}}\right] = c\times\mathrm{diag}\left\{\underbrace{0,1,\cdots,1}_{r+m},\underbrace{0,1,\cdots,1}_{q},\underbrace{0,1,\cdots,1}_{p}\right\} \tag{59}$$

其中, c 为一较大的正常数.

3) 给定多项式函数阶数 p, q. 根据前文所述进行适当调整.

4) 设定 R_t^2 的计算方法, 根据式 (54) 计算.

4　仿真实验

本节使用两组仿真实验验证该方法的有效性. 仿真 1 对参数初始值大幅度偏离参数真实值时算法的辨识结果进行了实验. 这是一个极端情况下的仿真实验, 用以验证算法在大初始偏差下的收敛性. 仿真 2 使用该算法对带有任意形式的输入输出非线性 H-W 系统进行辨识, 在较强干扰存在的情况下与 SVD 法进行比较.

4.1　仿真 1

待辨识 H-W 系统各部分特性如下列函数表示.

1) 输入非线性函数:

$$v = f_1(u) = u + 0.6u^2 - 0.6u^3$$

2) 线性部分传递函数:

$$G(z^{-1}) = \frac{z^{-1}+0.6z^{-2}+1.1z^{-3}}{1+0.7z^{-1}+0.4z^{-2}-0.4z^{-3}}$$

3) 输出非线性逆函数:

$$w = f_2^{-1}(y) = y - 0.5y^2 + 0.2y^3$$

系统输入为均值为 0、方差为 1 的白噪声序列, 同时加入均值为 0、方差为 10^{-4} 的白噪声干扰, 噪信比为 2 %.

为了验证算法的强收敛性, 设置参数初始值大幅度偏离参数真值, 令:

$$\hat{\boldsymbol{\theta}}_0 = (\underbrace{1,20,\cdots,20}_{r+m},\underbrace{1,20,\cdots,20}_{q},\underbrace{1,20,\cdots,20}_{p})^{\mathrm{T}}$$

选取初始参数误差协方差期望矩阵为

$$\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_0\tilde{\boldsymbol{\theta}}_0^{\mathrm{T}}\right] = 10^6\times\mathrm{diag}\left\{\underbrace{0,1,\cdots,1}_{r+m},\underbrace{0,1,\cdots,1}_{q},\underbrace{0,1,\cdots,1}_{p}\right\}$$

多项式函数阶数 p, q 与原系统相同, R_t^2 按下式计算: $R_t^2 = 10\boldsymbol{f}_{t-1}^{\mathrm{T}}\mathrm{E}\left[\tilde{\boldsymbol{\theta}}_{t-1}\tilde{\boldsymbol{\theta}}_{t-1}^{\mathrm{T}}\right]\boldsymbol{f}_{t-1}+0.0001$. 通过 1 000 步的仿真实验, 得到的参数误差平方和如图 2 所示.

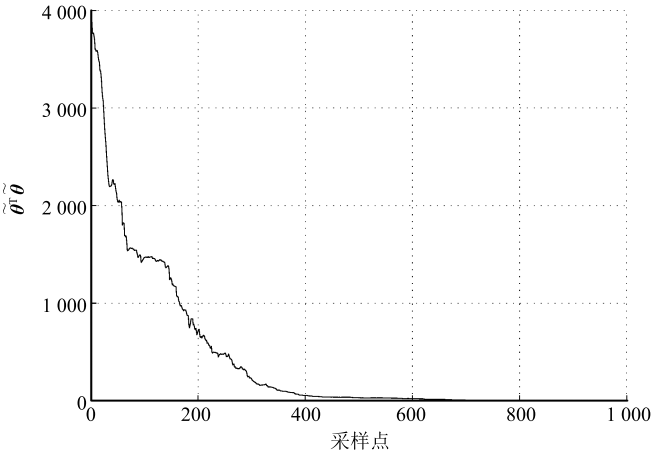


图 2　参数误差平方和

Fig. 2　Sum of square on parameter errors

由图 2 可见, 虽然参数初始值大幅度偏离真值, 但参数估计依然实现了一致收敛.

4.2　仿真 2

为了验证算法在较强干扰作用下对任意 H-W 系统的辨识能力, 本次仿真加入均值为 0、方差为 10^{-4} 的白噪声干扰, 噪信比为 7 %; 同时待辨识 H-W 系统输入输出非线性采用非多项式形式.

1) 输入非线性函数:

$$v = f_1(u) = \frac{u}{\sqrt{2.5-3u+3.5u^2}}$$

2) 线性部分传递函数:

$$G(z) = \frac{0.3z^{-1}+0.21z^{-2}}{1-0.8z^{-1}+0.4z^{-2}}$$

3) 输出非线性逆函数:

$$y = f_2(w) = \mathrm{e}^{1.5w} - 1$$

参数初始值的选取与第 3.2 节所给方法相同, 令: $\hat{\boldsymbol{\theta}}_0 = (\underbrace{1,0,\cdots,0}_{r+m},\underbrace{1,0,\cdots,0}_{q},\underbrace{1,0,\cdots,0}_{p})^{\mathrm{T}}$. p 和 q 分别取 5 和 6, 其他设定与仿真 1 相同. 使用本文方法与 SVD 进行比较, 经过 500 步的仿真实验并进行相应参数的增益调整后, 辨识结果如图 3 和图 4 所示. 辨识结束时, 线性部分参数估计误差以及非线性部分曲线拟合标准差如表 1 所示.

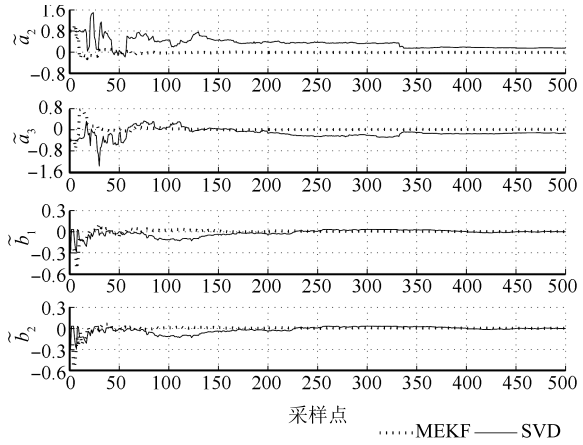


图 3 线性部分参数辨识曲线

Fig. 3 Parameter estimators of linear part

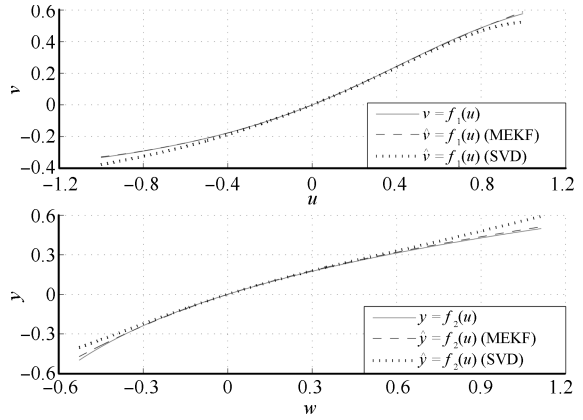


图 4 非线性部分辨识结果

Fig. 4 Curves of estimation on non-linear parts

表 1 辨识结果比较

Table 1 Comparison of estimation results with two methods

	MEKF	SVD
$\tilde{a}_2$	0.003200	0.153804
$\tilde{a}_3$	−0.005168	−0.130897
$\tilde{b}_1$	−0.000477	−0.009544
$\tilde{b}_2$	0.003712	−0.006466
输入非线性拟合标准差	0.002147	0.020388
输出非线性拟合标准差	0.007961	0.035925

由图 3 可见, 辨识过程中, SVD 受扰动的作用, 线性部分参数在辨识过程中出现较大波动, 偏离真值, 而 MEKF 受扰动影响较小, 由表 1 可见 MEKF 的辨识误差远低于 SVD. 对于非线性特性的辨识, 由图 4 以及表 1 可见, MEKF 的辨识精度高于 SVD.

5 结论

本文根据 EKF 的推导思路提出了一种递归算法, 用于在线辨识存在过程噪声的 H-W 系统. 该方法对 H-W 系统进行了严格的参数化; 在推导参数估计的递归更新公式过程中, 以参数误差平方和最小为目标函数, 有效避免了扰动对辨识结果的影响; 通过分析证明给出了算法一致收敛的条件以及重要参数 R_t^2 的设定方法, 扩大了算法的收敛域, 使算法得到更加广泛的适用范围. 从理论分析与仿真实验可见, 与两阶段 SVD 法相比, 本文提出的 MEKF 可以更准确地辨识 H-W 系统, 当扰动作用于系统时具有更高的辨识精度. 需要指出的是本文假设过程噪声为白噪声. 而在实际过程中, 过程噪声为有色噪声, 关于有色噪声的辨识可以参考文献 [24−26]. 因此本文研究的是一种理想状态下的 H-W 系统, 对于含有有色噪声 H-W 系统的辨识是今后这类系统辨识研究的方向. 此外, 该辨识算法需要输出非线性部分可逆, 这一条件限制了该辨识算法的应用, 如何去掉该假设也是需要进一步讨论的问题.

References

1 Li Yan, Mao Zhi-Zhong, Wang Yan, Yuan Ping, Jia Ming-Xing. Predictive control of Hammerstein-Wiener nonlinearity based on polytopic terminal region. *Acta Automatica Sinica*, 2011, **37**(5): 629−638
(李妍, 毛志忠, 王琰, 袁平, 贾明兴. 基于多面体终端域的 Hammerstein-Wiener 非线性预测控制. 自动化学报, 2011, **37**(5): 629−638)

2 Bai E W, Li K. Convergence of the iterative algorithm for a general Hammerstein system identification. *Automatica*, 2010, **46**(11): 1891−1896

3 Ding F, Liu X P, Liu G J. Identification methods for Hammerstein nonlinear systems. *Digital Signal Processing*, 2011, **21**(2): 215−238

4 Chen X, Fang H T. Recursive identification for Hammerstein systems with state-space model. *Acta Automatica Sinica*, 2010, **36**(10): 1460−1467

5 Zhao W X, Chen H F. Adaptive tracking and recursive identification for Hammerstein systems. *Automatica*, 2009, **45**(12): 2773−2783

6 Cai Z J, Bai E W. Making parametric Hammerstein system identification a linear problem. *Automatica*, 2011, **47**(8): 1806−1812

7 Bai E W, Cai Z J, Dudley-Javorosk S, Shields R K. Identification of a modified Wiener-Hammerstein system and its application in electrically stimulated paralyzed skeletal muscle modeling. *Automatica*, 2009, **45**(3): 736−743

8 Kobayashi Y, Shiotani Y, Hikita S, Nomura K. Identification of time-varying Wiener systems with unknown parameters. *Electronics and Communications in Japan*, 2010, **93**(4): 1−9

9 Biagiola S I, Figueroa J L. Wiener and Hammerstein uncertain models identification. *Mathematics and Computers in Simulation*, 2009, **79**(11): 3296−3313

10 Vörös J. Parameter identification of Wiener systems with multisegment piecewise-linear nonlinearities. *Systems and Control Letters*, 2007, **56**(2): 99−105

11 Bai E W. A blind approach to the Hammerstein-Wiener model identification. *Automatica*, 2002, **38**(6): 967−979

12 Bai E W. An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems. *Automatica*, 1998, **34**(3): 333−338

13 Li Yan, Mao Zhi-Zhong, Wang Yan, Yuan Ping, Jia Ming-Xing. Identification of Hammerstein-Wiener models based on bias compensation recursive least squares. *Acta Automatica Sinica*, 2010, **36**(1): 163−168
(李妍, 毛志忠, 王琰, 袁平, 贾明兴. 基于偏差补偿递推最小二乘的 Hammerstein-Wiener 模型辨识. 自动化学报, 2010, **36**(1): 163−168)

14 Gui Wei-Hua, Song Hai-Ying, Yang Chun-Hua. Hammerstein-Wiener model identified by least-squares-support-vector machine and its application. *Control Theory and Applications*, 2008, **25**(3): 393−397
(桂卫华, 宋海鹰, 阳春华. Hammerstein-Wiener 模型最小二乘向量机辨识及其应用. 控制理论与应用, 2008, **25**(3): 393−397)

15 Wang D Q, Ding F. Extended stochastic gradient identification algorithms for Hammerstein-Wiener ARMAX systems. *Computers and Mathematics with Applications*, 2008, **56**(12): 3157−3164

16 Zhu Y C. Estimation of an N-L-N Hammerstein-Wiener model. *Automatica*, 2002, **38**(9): 1607−1614

17 Macarthur J W. A new approach for nonlinear process identification using orthonormal bases and ordinal splines. *Journal of Process Control*, 2012, **22**(2): 375−389

18 Chen Y C, Wang J S. A Hammerstein-Wiener recurrent neural network with frequency-domain eigensystem realization algorithm for unknown system identification. *Journal of Universal Computer Science*, 2009, **15**(13): 2547−2565

19 Nordsjo A E, Zetterberg L H. Identification of certain time-varying nonlinear wiener and Hammerstein systems. *IEEE Transactions on Signal Processing*, 2001, **49**(3): 577−592

20 Boutayeb M, Aubry D. A strong tracking extended Kalman observer for nonlinear discrete-time systems. *IEEE Transactions on Automatic Control*, 1999, **44**(8): 1550−1556

21 Reif K, Gunther S, Yaz E, Unbehauen R. Stochastic stability of the discrete-time extended Kalman filter. *IEEE Transactions on Automatic Control*, 1999, **44**(4): 714−728

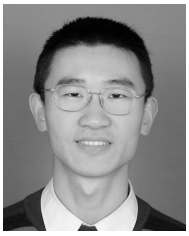
22 Reif K, Sonnemann F, Unbehauen R. An EKF-based nonlinear observer with a prescribed degree of stability. *Automatica*, 1998, **34**(9): 1119−1123

23 Boutayeb M, Rafaralahy H, Darouach M. Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems. *IEEE Transactions on Automatic Control*, 1997, **42**(4): 581−586

24 Zhang Yong, Yang Hui-Zhong. Bias compensation recursive least squares identification for output error systems with colored noises. *Acta Automatica Sinica*, 2007, **33**(10): 1053−1060
(张勇, 杨慧中. 有色噪声干扰输出误差系统的偏差补偿递推最小二乘辨识方法. 自动化学报, 2007, **33**(10): 1053−1060)

25 Pan Tian-Hong, Xue Zhen-Kuang, Li Shao-Yuan. An online multi-model identification algorithm based on subtractive clustering. *Acta Automatica Sinica*, 2009, **35**(2): 220−224
(潘天红, 薛振框, 李少远. 基于减法聚类的多模型在线辨识算法. 自动化学报, 2009, **35**(2): 220−224)

26 Ni Bo-Yi, Xiao De-Yun. A recursive identification method for non-uniformly sampled systems. *Acta Automatica Sinica*, 2009, **35**(12): 1520−1527
(倪博溢, 萧德云. 非均匀采样系统的一种递推辨识方法. 自动化学报, 2009, **35**(12): 1520−1527)

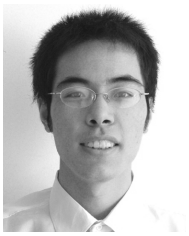


于 丰 东北大学博士研究生. 主要研究方向为非线性系统建模、优化与控制.
E-mail: yufeng3477@163.com
(**YU Feng** Ph. D. candidate at Northeastern University. His research interest covers nonlinear system model, optimization and control.)

毛志忠 东北大学教授. 主要研究方向为复杂系统建模、优化与控制. 本文通信作者.
E-mail: maozhizhong@ise.neu.edu.cn
(**MAO Zhi-Zhong** Professor at Northeastern University. His research interest covers complex system modeling, optimization and control. Corresponding author of this paper.)

贾明兴 东北大学教授. 主要研究方向为控制系统故障诊断和智能控制.
E-mail: jiamingxing@ise.neu.edu.cn
(**JIA Ming-Xing** Professor at Northeastern University. His research interest covers fault detection and intelligent control.)

袁 平 东北大学副教授. 主要研究方向为建模、软测量及智能控制.
E-mail: yuanping1@ise.neu.edu.cn
(**YUAN Ping** Associate professor at Northeastern University. His research interest covers modeling, soft measurement, and intelligent control.)



杨飞生 西北工业大学副教授. 主要研究方向为基于网络的系统, 时滞系统的容错控制, 参数辨识与智能自适应控制.
E-mail: fashion-young@hotmail.com
(**YANG Fei-Sheng** Associate professor at Northwestern Polytechnical University. His research interest covers network-based systems, fault-tolerant control for time-delay systems, parameter identification, and intelligent adaptive control.)