

• 计算机科学与技术 •

DOI:10.12454/j.jsuese.202301029



本刊网刊

未知模型参数下基于 Q 学习的卡尔曼状态估计算法

杨雯莹, 栾小丽*, 刘 飞

(江南大学 自动化研究所 轻工过程先进控制教育部重点实验室, 江苏 无锡 214122)

摘要:高精度的卡尔曼状态估计算法要求模型参数及噪声统计特性精确已知, 否则将导致估计性能显著下降, 甚至滤波发散。针对这一局限性, 在模型参数及噪声统计特性未知情况下, 提出一种边学习边估计的卡尔曼状态估计算法。算法采用 Q 学习策略迭代方法来解决模型信息未知的问题, 分为策略改进和策略评估两个部分。在策略改进阶段, 首先定义一个可评价状态估计值的状态动作价值函数, 即 Q 函数, 并进行公式变换, 使估计值仅依赖观测值而不依赖模型参数; 然后通过 Q 函数推导出获取系统状态估计值的估计策略。在策略评估阶段, 首先利用递推最小二乘法辨识 Q 函数的信息矩阵; 然后基于辨识到的信息矩阵, 遵循求得的估计策略, 执行相应动作, 更新状态变量的估计值; 最后将所提算法应用于估计二态多项式系统的状态与四联水箱系统的水位, 以此来验证算法的有效性与可行性, 并将所提算法与联合估计算法进行对比。仿真结果表明, 相较于联合估计算法, 当系统噪声为高斯噪声时, 所提算法针对两种系统状态的平均均方根误差分别降低了 34.66% 与 79.93%, 估计精度更高, 面对参数的不确定性表现出更强的鲁棒性; 当系统噪声为非高斯噪声时, 所提算法的估计精度与联合估计算法相近。此外, 在高斯噪声、非高斯噪声二态多项式系统与四联水箱系统的实验中, 算法运行时间相较于联合估计算法分别降低了 44.38%、45.03% 与 47.78%, 有效提高了算法的实时性。本文提出的估计算法能够为卡尔曼状态估计算法在实际工程应用和拓展中提供思路和方法。

关键词:卡尔曼状态估计; Q 学习; 未知模型参数; 未知噪声统计特性

中图分类号: TN713

文献标志码: A

文章编号: 2096-3246(2025)06-0335-09

在实际工业过程中, 受测量手段、成本等多方面限制, 系统状态变量的真实值往往无法直接获取, 需要通过一种成本低廉且运行可靠的状态估计方法来得到, 卡尔曼滤波(KF)算法应运而生^[1-2]。随着计算机技术的不断发展, KF 算法凭借其简单、高效且估计结果为全局最优等优势得以广泛研究与应用^[3-4], 并延伸出多种改进、拓展方法^[5-7]。然而, KF 算法的应用依赖于精确已知的系统模型参数与噪声统计特性, 且噪声需为高斯白噪声。这导致一旦出现模型失配或存在有色噪声的情况, KF 算法的估计性能将显著下降, 甚至引起滤波发散^[8]。自适应 KF 算法能够处理模型失配或存在有色噪声时的状态估计问题^[9], 常用的方法有 Sage-Husa 自适应 KF、基于新息的自适应 KF 与多模型 KF 等^[10-12]。

但是, 此类自适应估计算法大多只考虑了环境与真实系统发生偏离时的情况。而在实际应用过程中, 受系统复杂性、不确定性或者时间和资源的限制, 系统的模型参数和噪声统计特性往往难以详细获取, 甚至完全未知^[13-14]。

针对模型信息未知时的状态估计问题, 有两种主要解决方法。第一种解决方法是先辨识模型参数, 再估计内部状态^[15-16]。递归辨识^[17-18]和迭代辨识^[19-20]是两种重要的参数辨识方法。在最近的研究中, 研究者应用负梯度搜索和关键项分离技术对 Hammerstein 输出误差系统的参数辨识问题进行了研究, 提出了一种基于关键项分离的辅助模型递归梯度算法^[21]。Ding 等^[22]提出了一种滤波辅助模型分层广义扩展随机梯度辨识算法和一种滤波辅助模型分级多创新广义扩

收稿日期: 2023-12-15 修回日期: 2024-05-25 网络出版日期: 2024-06-03

基金项目: 国家自然科学基金项目(61991402; 61833007; 62503201)

作者简介: 杨雯莹(1999—), 女, 硕士生。研究方向: 控制工程及应用。E-mail: 6221913011@stu.jiangnan.edu.cn

* 通信作者: 栾小丽, 教授, E-mail: xlluan@jiangnan.edu.cn

展随机斜率辨识算法来辨识 Box-Jenkins 系统的参数。Yang 等^[23]针对非线性反馈系统的参数辨识问题,提出了一种基于层次梯度的迭代算法,以提高参数辨识精度。通过辨识算法得到精确的辨识参数,再结合状态估计算法,就能完成对系统内部状态的估计。另一种解决方法则是模型参数与状态的联合估计^[24-25]。在最近的研究中,Aslan 等^[26]提出了一种基于最大似然的粒子平滑期望最大化算法,联合估计血流动力学模型的状态和参数。Abolhasani 等^[27]通过将系统状态和未知参数组合为增强状态,并利用鲁棒正则化最小二乘法来处理不确定性,提出了一种增强状态鲁棒正则化最小二乘滤波器;然而,如果模型和测量的不确定性很大,该算法可能会表现出高度的保守性。Marcos 等^[28]采用随机的观点来减少保守性以解决此问题,提出了一种基于估计变化增加不确定性(EVIU)标准的类卡尔曼滤波器,称为 EVIU 滤波器,将未知参数视为状态增量,在模型参数未知的情况下进行状态估计。然而,上述研究均未考虑过程噪声与测量噪声统计特性未知的情况,且同时估计参数与状态将带来很大的计算负担。而时变且未知的噪声会对估计精度产生重大影响,针对噪声统计特性未知的情况,在最近的研究中,Wang 等^[29]提出了一种基于 Pearson type VII 分布的自适应滑动窗口异常鲁棒 KF 算法实现联合估计。但该算法仅考虑了部分参数含有不确定性的情况,且尚未解决算法计算复杂度高、运行时间长的问题。

为解决现有算法存在的问题,提出一种基于 Q 学习的卡尔曼滤波(Q-learning based KF, QL-KF)状态估计算法。利用 Q 学习算法的无模型特性解决模型参数未知的问题:首先,定义一个 Q 函数来评价状态估计值;接着,通过 Q 函数推导出获取状态估计值的估计策略;然后,使用递推最小二乘法辨识 Q 函数信息矩阵;最后,基于辨识结果,遵循估计策略,更新状态估计值,实现模型参数未知情况下的状态估计。

1 模型参数已知的状态估计算法

模型参数已知的状态估计算法以 KF 及其拓展算法为主流。以 KF 算法为例,针对线性离散系统建立的状态空间模型如式(1)所示,分为系统模型与测量模型。

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \boldsymbol{\omega}_k, \\ \mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k \end{cases} \quad (1)$$

式中: \mathbf{x}_k 为状态变量,下标 k 表示第 k 个时间步; \mathbf{y}_k 为观测变量; $\boldsymbol{\omega}_k$ 为过程噪声; \mathbf{v}_k 为测量噪声; \mathbf{A}_k 为状态转移矩阵; \mathbf{C}_k 为观测矩阵。过程噪声 $\boldsymbol{\omega}_k$ 与测量噪声 \mathbf{v}_k 均满足均值为 0 的正态分布,且相互独立,是两个高斯白噪

声序列,其统计特性如下。

$$\begin{cases} E(\boldsymbol{\omega}_k) = 0, E(\boldsymbol{\omega}_k \boldsymbol{\omega}_j^T) = \mathbf{q}_k \delta_{kj}; \\ E(\mathbf{v}_k) = 0, E(\mathbf{v}_k \mathbf{v}_j^T) = \mathbf{r}_k \delta_{kj}; \\ E(\boldsymbol{\omega}_k \mathbf{v}_j^T) = 0 \end{cases} \quad (2)$$

式中:下标 j 表示第 j 个时间步; $E(\cdot)$ 为 (\cdot) 的期望值; \mathbf{q}_k 为过程噪声 $\boldsymbol{\omega}_k$ 的协方差矩阵; \mathbf{r}_k 为测量噪声 \mathbf{v}_k 的协方差矩阵; δ_{kj} 为 Kronecker δ 函数,满足式(3):

$$\delta_{kj} = \begin{cases} 1, k=j; \\ 0, k \neq j \end{cases} \quad (3)$$

KF 算法使用 1 组递推公式,即式(4)~(8)得到最优估计^[30],并不断调整卡尔曼增益,使估计的误差协方差最小。

$$\hat{\mathbf{x}}_k^- = \mathbf{A}_k \hat{\mathbf{x}}_{k-1} \quad (4)$$

$$\mathbf{P}_k^- = \mathbf{A}_k \mathbf{P}_{k-1} \mathbf{A}_k^T + \mathbf{q}_k \quad (5)$$

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{r}_k)^{-1} \quad (6)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k^-) \quad (7)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{P}_k^- \quad (8)$$

式(4)~(8)中: $\hat{\mathbf{x}}_k^-$ 为先验估计值; $\hat{\mathbf{x}}_k$ 为后验估计值; \mathbf{P}_k^- 和 \mathbf{P}_k 分别为先验估计值和后验估计值与真实值之间的误差协方差矩阵; \mathbf{I} 为单位矩阵; \mathbf{K}_k 为卡尔曼增益, $\mathbf{K}_k \in [0, \mathbf{C}_k^{-1}]$ 。

由式(4)~(8)可知,后验估计由先验估计和观测获得,而先验估计取决于模型参数,且卡尔曼增益的不断调整依赖于两个噪声的协方差矩阵。因此,使用 KF 算法进行状态估计需要准确了解模型参数和噪声统计特性,否则将导致模型误差,引起滤波发散^[31]。因此,针对复杂的实际工业过程,不依赖系统模型参数与噪声统计特性的状态估计算法亟待研究。

2 模型参数未知的状态估计算法

模型参数未知的状态估计算法采用策略迭代方法。

根据最优状态估计的要求,针对线性离散系统,依据式(1)定义代价函数^[32]如下。

$$c_k = (\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_k^-)^T \mathbf{E}_k^{-1} (\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_k^-) + (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k)^T \mathbf{F}_k^{-1} (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k) \quad (9)$$

式中, c_k 为代价函数, \mathbf{E}_k 为 $(\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_k^-)$ 误差的协方差矩阵, \mathbf{F}_k 为 $(\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k)$ 误差的协方差矩阵。在算法的实际应用中, \mathbf{E}_k 和 \mathbf{F}_k 是两个可调的权重矩阵。

最小化代价函数可使估计值尽可能地同时接近先验值与观测值,与真实值之间的差异最小,实现最优估计。根据权重矩阵 \mathbf{E}_k 和 \mathbf{F}_k ,无须已知系统噪声统计特性,卡尔曼增益 \mathbf{K}_k 可以表示为:

$$\mathbf{K}_k = \mathbf{E}_k \mathbf{C}_k^T (\mathbf{C}_k \mathbf{E}_k \mathbf{C}_k^T + \mathbf{F}_k)^{-1} \quad (10)$$

在卡尔曼估计中,当前时刻的估计值只决定下一时刻的先验值,需与下一时刻的观测值共同完成估计。换言之,当前时刻的估计值只部分地影响下一时刻的估计值,但包含了先前所有时刻的信息。

基于此,可将二次型状态价值函数 V_k 表示为:

$$V_k = \sum_{i=0}^k \gamma^i c_{k-i} = \sum_{i=0}^k \gamma^i [(\hat{x}_{k-i} - \hat{x}_{k-i}^-)^T E_{k-i}^{-1} (\hat{x}_{k-i} - \hat{x}_{k-i}^-) + (\mathbf{y}_{k-i} - \mathbf{C}_{k-i} \hat{x}_{k-i})^T \mathbf{F}_{k-i}^{-1} (\mathbf{y}_{k-i} - \mathbf{C}_{k-i} \hat{x}_{k-i})] \quad (11)$$

式中, γ 为折扣因子, $\gamma \in [0, 1]$ 。

模型参数未知情况下,求解最优估计问题采用的策略迭代方法分为策略改进和策略评估两个部分。首先,使用 Q 函数生成一个改进的估计策略;接着,估计 Q 函数。两个步骤交替进行,实现状态估计值的增量更新。

根据贝尔曼最优性原理^[33],将 Q 函数定义为:

$$Q_k = c_k + \gamma V_{k-1} = c_k + \gamma Q_{k-1} \quad (12)$$

将式(9)和(11)代入式(12)得:

$$Q_k = c_k + \gamma V_{k-1} \approx (\hat{x}_k - \hat{x}_k^-)^T \mathbf{E}_k^{-1} (\hat{x}_k - \hat{x}_k^-) + (\mathbf{y}_k - \mathbf{C}_k \hat{x}_k)^T \mathbf{F}_k^{-1} (\mathbf{y}_k - \mathbf{C}_k \hat{x}_k) + \gamma [(\hat{x}_{k-1} - \hat{x}_{k-1}^-)^T \mathbf{E}_{k-1}^{-1} (\hat{x}_{k-1} - \hat{x}_{k-1}^-) + (\mathbf{y}_{k-1} - \mathbf{C}_{k-1} \hat{x}_{k-1})^T \mathbf{F}_{k-1}^{-1} (\mathbf{y}_{k-1} - \mathbf{C}_{k-1} \hat{x}_{k-1})] \quad (13)$$

由于距当前时刻越远的信息,对当前估计值的影响越小,式(12)省略了 $k-1$ 时刻前的观测值与估计值。

与联合估计算法不同,QL-KF算法的优化目标是找到能最小化累积代价,即最小化 Q 函数的估计策略,而非同时估计系统的状态与参数。通过取 Q 函数的偏导数,将其设置为0,就可得到 Q 函数取最小值时对应的最优估计值。相应的,获取系统状态估计值的最优策略为:

$$\hat{x}_k = \arg \min_{\hat{x}} Q_k \quad (14)$$

根据上述对KF算法的分析可知,状态估计值受先验值的影响,而先验值需要依据系统模型参数,通过式(4)求得。因此,在未知系统模型参数的情况下,可以根据式(7)进行公式变换,得到:

$$\hat{x}_k^- = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k)^{-1} \hat{x}_k - (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k)^{-1} \mathbf{K}_k \mathbf{y}_k \quad (15)$$

将先验值变换成不依赖模型参数而依赖观测值与估计值的形式,使得估计值的求取仅依赖观测值。

将式(15)代入 Q 函数,并将其展开,写成二次型:

$$Q_k = [\hat{x}_k - \alpha_k (\hat{x}_k - \mathbf{K}_k \mathbf{y}_k)]^T \mathbf{E}_k^{-1} [\hat{x}_k - \alpha_k (\hat{x}_k - \mathbf{K}_k \mathbf{y}_k)] + (\mathbf{y}_k - \mathbf{C}_k \hat{x}_k)^T \mathbf{F}_k^{-1} (\mathbf{y}_k - \mathbf{C}_k \hat{x}_k) + \gamma [(\mathbf{A}_k^{-1} \alpha_k (\hat{x}_k - \mathbf{K}_k \mathbf{y}_k) - \alpha_{k-1} (\hat{x}_{k-1} - \mathbf{K}_{k-1} \mathbf{y}_{k-1}))^T \cdot \mathbf{E}_{k-1}^{-1} [\mathbf{A}_k^{-1} \alpha_k (\hat{x}_k - \mathbf{K}_k \mathbf{y}_k) - \alpha_{k-1} (\hat{x}_{k-1} - \mathbf{K}_{k-1} \mathbf{y}_{k-1})] + [\mathbf{y}_{k-1} - \mathbf{C}_k \mathbf{A}_k^{-1} \alpha_k (\hat{x}_k - \mathbf{K}_k \mathbf{y}_k)]^T \cdot \mathbf{F}_{k-1}^{-1} [\mathbf{y}_{k-1} - \mathbf{C}_k \mathbf{A}_k^{-1} \alpha_k (\hat{x}_k - \mathbf{K}_k \mathbf{y}_k)]] = [\hat{x}_k, \hat{x}_{k-1}, \mathbf{y}_k, \mathbf{y}_{k-1}]^T \mathbf{H}_k [\hat{x}_k, \hat{x}_{k-1}, \mathbf{y}_k, \mathbf{y}_{k-1}] \quad (16)$$

其中,

$$\alpha_k = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k)^{-1} \quad (17)$$

$$\mathbf{H}_{11,k} = (\mathbf{I} - \alpha_k)^T \mathbf{E}_k^{-1} (\mathbf{I} - \alpha_k) + \mathbf{C}_k^T \mathbf{F}_k^{-1} \mathbf{C}_k + \gamma (\mathbf{A}_k^{-1} \alpha_k)^T \mathbf{E}_{k-1}^{-1} \mathbf{A}_k^{-1} \alpha_k + \gamma (\mathbf{C}_k \mathbf{A}_k^{-1} \alpha_k)^T \mathbf{F}_{k-1}^{-1} \mathbf{C}_k \mathbf{A}_k^{-1} \alpha_k \quad (18)$$

$$\mathbf{H}_{12,k} = -\gamma (\mathbf{A}_k^{-1} \alpha_k)^T \mathbf{E}_{k-1}^{-1} \alpha_{k-1} \quad (19)$$

$$\mathbf{H}_{13,k} = (\mathbf{I} - \alpha_k)^T \mathbf{E}_k^{-1} \alpha_k \mathbf{K}_k - \mathbf{C}_k^T \mathbf{F}_k^{-1} - \gamma (\mathbf{A}_k^{-1} \alpha_k)^T \mathbf{E}_{k-1}^{-1} \mathbf{A}_k^{-1} \alpha_k \mathbf{K}_k - \gamma (\mathbf{C}_k \mathbf{A}_k^{-1} \alpha_k)^T \mathbf{F}_{k-1}^{-1} \mathbf{C}_k \mathbf{A}_k^{-1} \alpha_k \mathbf{K}_k \quad (20)$$

$$\mathbf{H}_{14,k} = \gamma (\mathbf{A}_k^{-1} \alpha_k)^T \mathbf{E}_{k-1}^{-1} \alpha_{k-1} - \gamma (\mathbf{C}_k \mathbf{A}_k^{-1} \alpha_k)^T \mathbf{F}_{k-1}^{-1} \quad (21)$$

式(16)~(21)中: \mathbf{H}_k 为信息矩阵,是一个对称正定矩阵,包含了先验值的所有信息; $\mathbf{H}_{i,k}$ 为矩阵 \mathbf{H}_k 中第 i 行的第 i 个分块矩阵, $i=1,2,3,4$; α_k 为过程变量。

将式(16)~(21)代入式(14)可得估计策略:

$$\hat{x}_k = -(\mathbf{H}_{11,k})^{-1} (\mathbf{H}_{12,k} \hat{x}_{k-1} + \mathbf{H}_{13,k} \mathbf{y}_k + \mathbf{H}_{14,k} \mathbf{y}_{k-1}) \quad (22)$$

为根据估计策略获取状态估计值,需在策略评估阶段对 Q 函数进行估计。 Q 函数的估计与更新是 Q 学习的核心。在QL-KF算法中,每次更新只涉及当前状态、观测、代价与上一时刻状态,计算复杂度较低。

Q 函数的估计问题可以转化为 Q 函数信息矩阵 \mathbf{H}_k 的辨识问题,使用递推最小二乘法辨识信息矩阵 \mathbf{H}_k 。

首先,将二次型转换为以下形式:

$$\mathbf{x}^T \mathbf{L} \mathbf{x} = \bar{\mathbf{x}}^T \boldsymbol{\Theta}(\mathbf{L}) \quad (23)$$

式中: $\bar{\mathbf{x}}$ 是由向量 \mathbf{x} 各元素的所有二次基函数构成的向量,且 $\bar{\mathbf{x}}^T = [x_1^2, x_1 x_2, \dots, x_1 x_n, x_2^2, x_2 x_3, \dots, x_2 x_n, \dots, x_n^2]$, x_i 为向量 \mathbf{x} 的第 i 个分量, $i \in [1, n]$; 向量 $\boldsymbol{\Theta}(\mathbf{L})$ 中元素由矩阵 \mathbf{L} 的 n 个对角元素和 $\frac{n(n+1)}{2} - n$ 个不同双元素之和 ($L_{ij} + L_{ji}$) 组成。

相应地,根据式(12)和(16),得到代价函数在线辨识模型的一般形式:

$$c_k (\hat{x}_k, \hat{x}_{k-1}, \mathbf{y}_k, \mathbf{y}_{k-1}) z Q_k (\hat{x}_k, \hat{x}_{k-1}, \mathbf{y}_k, \mathbf{y}_{k-1}) - \gamma Q_{k-1} (\hat{x}_{k-1}, \hat{x}_{k-2}, \mathbf{y}_{k-1}, \mathbf{y}_{k-2}) = \overline{[\hat{x}_k, \hat{x}_{k-1}, \mathbf{y}_k, \mathbf{y}_{k-1}]}^T \boldsymbol{\Theta}(\mathbf{H}_k) - \gamma \overline{[\hat{x}_{k-1}, \hat{x}_{k-2}, \mathbf{y}_{k-1}, \mathbf{y}_{k-2}]}^T \boldsymbol{\Theta}(\mathbf{H}_k) = \phi_k^T \theta_k \quad (24)$$

式中: $\overline{[\hat{x}_k, \hat{x}_{k-1}, \mathbf{y}_k, \mathbf{y}_{k-1}]}$ 为向量 $\hat{x}_k, \hat{x}_{k-1}, \mathbf{y}_k, \mathbf{y}_{k-1}$ 中各元素的所有二次基函数构成的向量; ϕ_k 为信息向量, $\phi_k = \overline{[\hat{x}_k, \hat{x}_{k-1}, \mathbf{y}_k, \mathbf{y}_{k-1}]} - \gamma \overline{[\hat{x}_{k-1}, \hat{x}_{k-2}, \mathbf{y}_{k-1}, \mathbf{y}_{k-2}]}$, 由已知信息构成; θ_k 为待辨识的参数向量, $\theta_k = \boldsymbol{\Theta}(\mathbf{H}_k)$, 包含所有未知信息,需使用以下递推最小二乘法公式进行辨识。

$$\begin{cases} e_k(s) = c_k - \phi_k^T \hat{\theta}_k(s-1), \\ \hat{\theta}_k(s) = \hat{\theta}_k(s-1) + \mathbf{P}_k(s-1) \phi_k e_k(s) [1 + \phi_k^T \mathbf{P}_k(s-1) \phi_k]^{-1}, \\ \mathbf{P}_k(s) = \mathbf{P}_k(s-1) - \frac{\mathbf{P}_k(s-1) \phi_k \phi_k^T \mathbf{P}_k(s-1)}{1 + \phi_k^T \mathbf{P}_k(s-1) \phi_k}, \\ \mathbf{P}_k(0) = \mathbf{P}_0 \end{cases} \quad (25)$$

式中: \mathbf{e}_k 为代价函数的误差; $\hat{\boldsymbol{\theta}}_k(s)$ 为 $\boldsymbol{\theta}_k$ 的第 s 个估计值; \mathbf{P}_0 为 \mathbf{P}_k 的初始值, $\mathbf{P}_0 = \beta \mathbf{I}$, β 是一个很大的正常数; s 为辨识 \mathbf{H}_k 的时间步数。

利用辨识出的 \mathbf{H}_k , 根据最优策略, 执行以下操作:

$$\hat{\mathbf{x}}_{k+1} = -(\mathbf{H}_{11,k})^{-1}(\mathbf{H}_{12,k}\hat{\mathbf{x}}_k + \mathbf{H}_{13,k}\mathbf{y}_{k+1} + \mathbf{H}_{14,k}\mathbf{y}_k) \quad (26)$$
便可得到下一时刻状态的最优估计值。

综上所述, 在得到一系列系统观测值的情况下, 无须使用系统模型参数, 也无须已知系统噪声统计特性, 只要赋予状态估计值及卡尔曼增益一些随机初值, 就可以实现状态估计。

3 仿真实验

为验证本文所提估计算法在模型参数与噪声统计特性未知的情况下的有效性与可行性, 将算法运用于估计二态多项式系统的状态和四联水箱系统的水位, 给出相应仿真结果。使用均方根误差 (RMSE, 记为 E_{RMSE}) 与平均均方根误差 (ARMSE, 记为 E_{ARMSE}) 作为性能评价指标。 k 时刻状态估计的 RMSE 为:

$$E_{\text{RMSE}}(k) = \sqrt{\frac{1}{M} \sum_{i=1}^M (x_{k,j}^i - \hat{x}_{k,j}^i)^2} \quad (27)$$

式中, M 为蒙特卡罗模拟次数, $x_{k,j}^i$ 为 \mathbf{x}_k 在第 i 个时间步的第 j 个分量的值, j 代表状态变量的个数, 即 $\mathbf{x}_k = [x_{k,1}, x_{k,2}, \dots, x_{k,j}]$ 。

ARMSE 定义为:

$$E_{\text{ARMSE}} = \frac{1}{N} \sum_{i=1}^N E_{\text{RMSE}}(i) \quad (28)$$

式中, N 为 RMSE 的个数。

3.1 二态多项式系统

3.1.1 高斯噪声

将 QL-KF 算法应用于二态多项式系统:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \boldsymbol{\omega}_k, \boldsymbol{\omega}_k \sim \mathcal{N}(0, \mathbf{q}_k); \\ \mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k, \mathbf{v}_k \sim \mathcal{N}(0, \mathbf{r}_k) \end{cases} \quad (29)$$

式中, $\mathbf{A}_k = \begin{bmatrix} 1.618 & 1 \\ -0.618 & 0 \end{bmatrix}$, $\mathbf{C}_k = [1 \ 0]$ 。

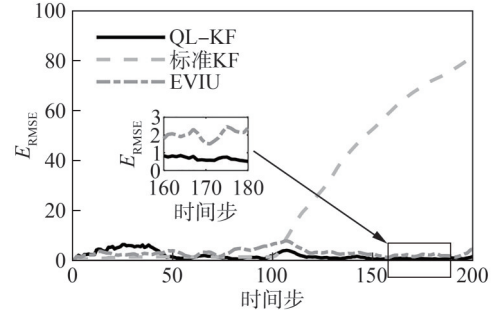
假设过程噪声与测量噪声皆为零均值的高斯白噪声, 且 $\mathbf{q}_k = \mathbf{I}$, $\mathbf{r}_k = \mathbf{I}^2$ 。为表征 QL-KF 算法, 对系统模型参数施加不确定性。表 1 为二项式多态系统的仿真环境。为增加模拟的可信度, 引入蒙特卡罗实验, 设置模拟次数 $M=30$, 并将本文算法与标准 KF 算法、以 EVIU 算法^[28] 为例的参数状态联合估计算法进行对比。

表 1 二态多项式系统的仿真环境

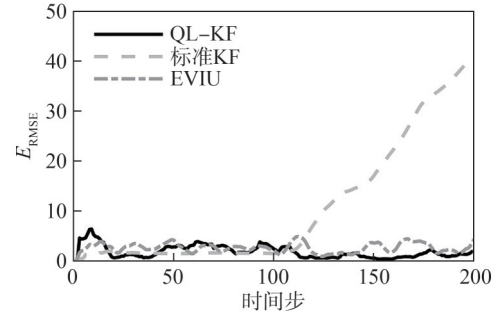
Tab. 1 simulation environment for the two-state polynomial system

变化参数	$\Delta \mathbf{A}_k$	$\Delta \mathbf{q}_k$	$\Delta \mathbf{r}_k$
变化值	$-0.9 \mathbf{A}_k$	$-0.9 \mathbf{q}_k$	$-0.9 \mathbf{r}_k$

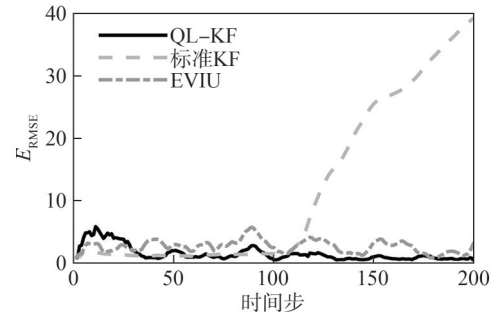
由于式(24)包含 $\hat{\mathbf{x}}_k$ 、 $\hat{\mathbf{x}}_{k-1}$ 、 $\hat{\mathbf{x}}_{k-2}$ 3 个时刻的状态估计值, 因此, 算法从时间步 $k=3$ 开始, 并对 $\hat{\mathbf{x}}_1$ 、 $\hat{\mathbf{x}}_2$ 、 $\hat{\mathbf{x}}_3$ 赋予随机初值。QL-KF 算法流程始于 $\mathbf{x}_1 = [0.3, 0.3]$, $\mathbf{x}_2 = [0.5, 0.5]$, $\mathbf{x}_3 = [1, 1]$, 并对 200 个时间步内的状态进行估计。图 1 为高斯噪声二态多项式系统不同参数变化时各算法的 RMSE。表 2 为高斯噪声下 3 种算法性能对比。



(a) $\Delta \mathbf{A}_k = -0.9 \mathbf{A}_k$ 时各算法的 RMSE



(b) $\Delta \mathbf{q}_k = -0.9 \mathbf{q}_k$ 时各算法的 RMSE



(c) $\Delta \mathbf{r}_k = -0.9 \mathbf{r}_k$ 时各算法的 RMSE

图 1 高斯噪声二态多项式系统不同参数变化时各算法的 RMSE

Fig. 1 RMSE of each algorithm when different parameters change for the two-state polynomial system with Gaussian noise

表 2 高斯噪声下 3 种算法性能对比

Tab. 2 Performance comparison of three algorithms under Gaussian noise

估计算法	E_{ARMSE}			平均耗时/s
	$\Delta \mathbf{A}_k$	$\Delta \mathbf{q}_k$	$\Delta \mathbf{r}_k$	
QL-KF	1.996 2	1.876 4	1.540 4	0.732
EVIU	3.033 2	2.625 1	2.622 6	1.316
标准KF	50.455 6	22.056 5	22.477 4	0.782

由图 1 和表 1、2 可见, 面对时变且未知的各模型参数, QL-KF 算法的 RSME 表现出良好的收敛趋势, 体现了算法的有效性。随机给定的初始估计值导致

QL-KF 算法初始时刻的 RMSE 可能略大。并且,由于 Q 学习算法在应用过程中需要一定的数据累积,当估计时间步较小时, RMSE 略有波动,但随着迭代次数的增加呈下降趋势,并逐渐趋于稳定。在第 101 个时间步,由于参数的变化, QL-KF 算法的 RMSE 有所增加,但增幅较小且快速恢复到较低水平,最终稳定在特定值附近,显示出较强的鲁棒性。

对比标准 KF 算法与 QL-KF 算法,标准 KF 算法在已知模型参数时, RMSE 较小且十分平稳。但面对模型参数改变时,在两者均未知模型参数的情况下,本文提出的 QL-KF 算法估计精度明显优于标准 KF 算法,且具有更强的鲁棒性。

与同样未知各模型参数的 EVIU 算法相比,在前 100 个时间步内,两者估计精度相当。在第 101 个时间步,面对各参数的变化, EVIU 算法的 RMSE 增幅更大。并且,随着迭代次数的增加, QL-KF 算法稳定后的 RMSE 更小且更为平稳。此外,由于 QL-KF 算法在每个时间步只需更新一次 Q 函数与估计策略,且为增量更新,算法计算量较小;而 EVIU 算法在每个时间步都要对新的观测数据进行全面处理,同时更新系统状态和参数估计,需要在每次迭代中进行大量的计算。

因此,与 EVIU 算法相比, QL-KF 算法的运行速度更快。综上所述,本文提出的 QL-KF 算法的 ARMSE 相较于 EVIU 算法平均降低约 34.66%, 估计精度更高;面对参数的不确定性时表现出更强的鲁棒性;算法平均耗时较少,与标准 KF 算法相当,相较于 EVIU 算法降低约 44.38%, 实时性高。

3.1.2 非高斯噪声

以均匀噪声为例,观察 QL-KF 算法针对非高斯噪声系统的估计性能。

假设二态多项式系统的过程噪声与测量噪声皆为均值为 $(0+10)/2$ 、方差为 $(10-0)^2/12$ 的均匀噪声。实验仍使用表 1 中的仿真环境,设置蒙特卡罗模拟次数 $M=30$ 。QL-KF 算法流程始于 $\mathbf{x}_1=[0.3, 0.3], \mathbf{x}_2=[0.5, 0.5], \mathbf{x}_3=[1, 1]$, 并对 200 个时间步内的状态进行估计。图 2 为非高斯噪声二态多项式系统不同参数变化时各算法的 RMSE。表 3 为非高斯噪声下 3 种算法性能对比结果。

由图 2 和表 1、3 可见, QL-KF 算法对均匀噪声具有一定的鲁棒性, RMSE 仍呈现收敛趋势。基于一定量的数据累积,估计器能够逐渐调整对状态的估计,使得误差逐步减小,并在长期运行后达到稳态。但非高斯噪声会导致 QL-KF 算法的估计误差增大,不同参数变化下 ARMSE 平均约为 14.28。算法的估计精度高于标准 KF 算法,与 EVIU 算法相当。平均耗时相较于 EVIU 算法降低约 45.03%, 仍具有较高的实时性。

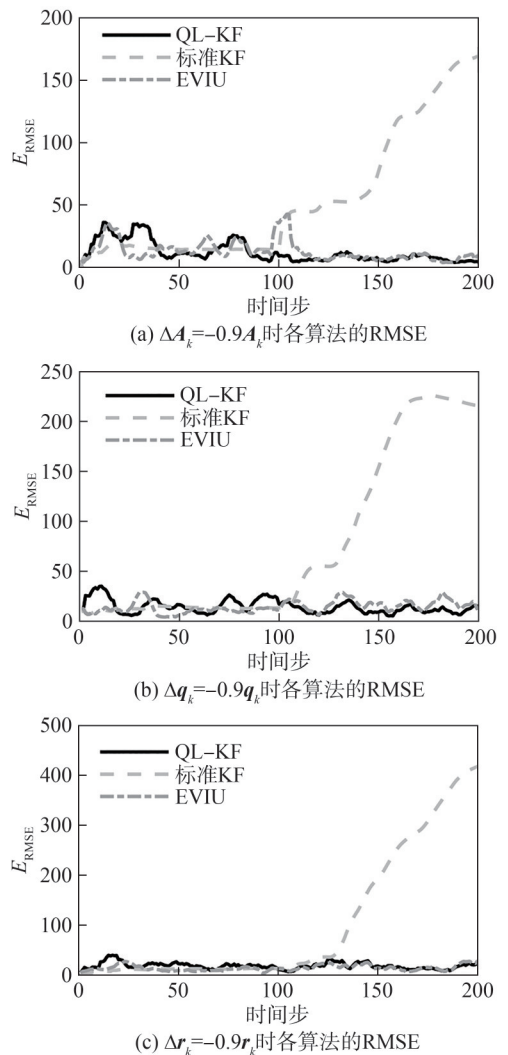


图 2 非高斯噪声二态多项式系统不同参数变化时各算法的 RMSE

Fig. 2 RMSE of each algorithm when different parameters change for the two-state polynomial system with non-Gaussian noise

表 3 非高斯噪声下 3 种算法性能对比

Tab. 3 Performance comparison of three algorithms under non-Gaussian noise

估计算法	E_{ARMSE}			平均耗时/s
	ΔA_k	Δq_k	Δr_k	
QL-KF	12.463 3	14.875 9	15.513 5	0.713
EVIU	12.981 7	14.130 5	14.186 2	1.297
标准 KF	90.730 4	169.691 9	262.512 5	0.778

3.2 四联水箱系统

本节使用四联水箱系统^[34]来验证算法的有效性。图 3 为四联水箱系统概述图。由图 3 可知,两台水泵使用两股分流将水送入不同的水箱,水箱 3 和 4 也分别将水送入水箱 1 和 2。上部水箱中的水只能排入其下方的水箱,从水箱底部排出的水直接流入大型蓄水池。系统状态向量定义如下。

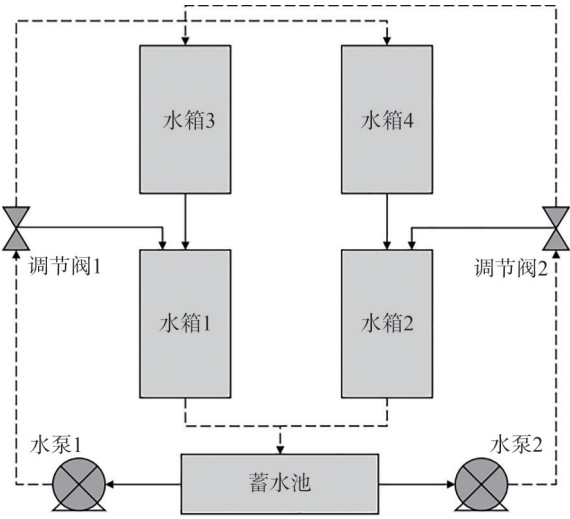


图3 四联水箱系统概述图

Fig. 3 Profile of the quadruple water tank system

$$\mathbf{x}_k = [h_1(k) - s_1, h_2(k) - s_2, h_3(k) - s_3, h_4(k) - s_4]^T \quad (30)$$

式中： $h_i(k)$ 为水箱*i*的水位， $i \in [1, 4]$ ； s_i 为水箱水位的设定值， $i \in [1, 4]$ ， $s_1 = 0.1226 \text{ m}$ ， $s_2 = 0.2602 \text{ m}$ ， $s_3 = 0.0101 \text{ m}$ ， $s_4 = 0.0408 \text{ m}$ 。

于是，四联水箱系统的状态空间模型可用式(1)描述，参数为：

$$A_k = \begin{bmatrix} 0.9321 & 0 & 0.4173 & 0 \\ 0 & 0.9395 & 0 & 0.1123 \\ 0 & 0 & 0.8613 & 0 \\ 0 & 0 & 0 & 0.9090 \end{bmatrix},$$

$$C_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

假设过程噪声与测量噪声皆为0均值的高斯白噪声，且

$$q_k = \text{diag}(10^{-4}, 10^{-4}, 10^{-4}, 10^{-4}), r_k = I.$$

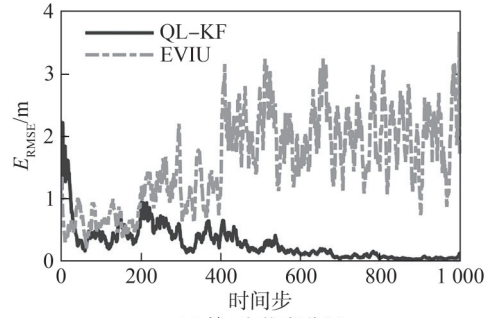
对系统模型参数施加不确定性。表4为四联水箱系统的仿真环境。设置蒙特卡罗模拟次数 $M = 30$ 。算法流程始于 $\mathbf{x}_1 = [0.3, 0.3, 0.3, 0.3]$ ， $\mathbf{x}_2 = [0.5, 0.5, 0.5, 0.5]$ ， $\mathbf{x}_3 = [1, 1, 1, 1]$ ，并对1000个时间步内的状态进行估计。图4为2种算法针对四联水箱系统的RMSE。表5为四联水箱系统两种算法性能对比结果。

由图4与表4、5可见，QL-KF算法针对两个状态分量的RMSE均表现出良好的趋势，验证了算法有效性。具体而言，在初始时刻，QL-KF的RSME较大，但此后便迅速收敛；在后续时间步中，参数的不确定性使得RMSE有所增加，但随后快速减小，呈收敛趋势，并随着参数的稳定逐渐趋于稳态。而EVIU算法受参数不确定性影响较大。与其相比，QL-KF算法鲁棒性更强，稳定后具有更高的精度，且RMSE更加平稳，ARMSE平均减小79.93%，算法平均耗时降低47.78%，有效提高了实时性。

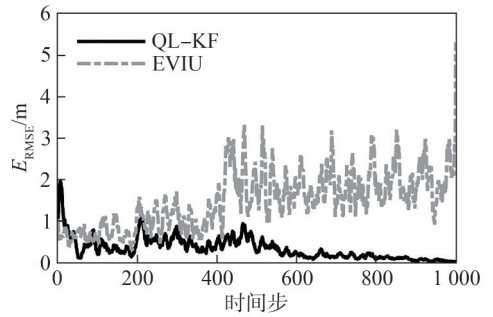
表4 四联水箱系统的仿真环境

Tab. 4 Simulation environment for the quadruple water tank system

时间步	ΔA_k	Δr_k	Δq_k
1~200	0	0	0
201~400	0	$2r_k$	0
401~600	$-0.93A_k$	$2r_k$	$-0.5q_k$
601~1000	0	0	0



(a) 第1个状态分量



(b) 第2个状态分量

图4 2种算法针对四联水箱系统的RMSE

Fig. 4 RMSE of two algorithms for the quadruple water tank system

表5 四联水箱系统2种算法性能对比

Tab. 5 Performance comparison of two algorithms in the quadruple water tank system

估计算法	E_{ARMSE}		平均耗时/s
	第1个状态分量	第2个状态分量	
QL-KF	0.2800	0.3386	1.731
EVIU	1.5965	1.4976	3.315

4 结论

本文针对使用卡尔曼状态估计算法需要系统模型参数及系统噪声统计特性精确已知这一局限性，提出了一种基于Q学习的卡尔曼状态估计算法。此算法仅赋予状态一些随机的初始估计值，就能够在系统模型参数及噪声统计特性未知的情况下，实现状态估计。仿真结果表明，本文提出的QL-KF算法能够在未知系统模型参数及噪声统计特性的情况下，无须辨识系统参数，仅利用观测值得到系统内部状态的估计。

算法的估计性能受系统噪声类型的影响。面对高斯噪声系统时,本文算法估计精度高、鲁棒性强、实时性高。但面对非高斯噪声系统时,估计精度有所下降,未来将考虑进一步提升估计精度。

参考文献:

- [1] Kordestani M, Safavi A A, Saif M. Recent survey of large-scale systems: Architectures, controller strategies, and industrial applications[J]. *IEEE Systems Journal*, 2021, 15(4): 5440–5453.
- [2] Tabacek J, Havlena V. Reduction of prediction error sensitivity to parameters in Kalman filter[J]. *Journal of the Franklin Institute*, 2022, 359(3): 1303–1326.
- [3] Chen Yifan, Han Haiqian, Zhang Yi, et al. Dynamic inversion of hydrodynamic parameters of plain river network [J]. *Advanced Engineering Sciences*, 2019, 51(2): 13–20. [陈一帆, 韩海骞, 张翼, 等. 城镇化平原河网水动力参数动态反演[J]. *工程科学与技术*, 2019, 51(2): 13–20.]
- [4] Li Qinwen, Wang Zhiqian, Wang Wenrui, et al. A model predictive obstacle avoidance method based on dynamic motion primitives and a Kalman filter[J]. *Asian Journal of Control*, 2023, 25(2): 1510–1525.
- [5] Miao Kelei, Zhang Wenan, Qiu Xiang. An adaptive unscented Kalman filter approach to secure state estimation for wireless sensor networks[J]. *Asian Journal of Control*, 2023, 25(1): 629–636.
- [6] Park G. Optimal vehicle position estimation using adaptive unscented Kalman filter based on sensor fusion[J]. *Mechanics*, 2024, 99: 103144.
- [7] Lu Xin, Liu Zhong, Zhang Hongxin, et al. A Gaussian nonlinear iterated update filter[J]. *Advanced Engineering Sciences*, 2017, 49(4): 111–118. [陆欣, 刘忠, 张宏欣, 等. 一种高斯型非线性迭代更新滤波器[J]. *工程科学与技术*, 2017, 49(4): 111–118.]
- [8] Pan Cheng, Gao Jingxiang, Li Zengke, et al. Multiple fading factors-based strong tracking variational Bayesian adaptive Kalman filter[J]. *Measurement*, 2021, 176: 109139.
- [9] Zeng Jingcong, Shi Yuanfeng, Dai Kaoshan, et al. Real-time structural displacement estimation by fusing acceleration and displacement data with adaptive Kalman filter[J]. *Advanced Engineering Sciences*, 2023, 55(4): 188–196. [曾竞聪, 施袁锋, 戴靠山, 等. 基于自适应卡尔曼滤波加速度与位移融合的结构位移实时估计[J]. *工程科学与技术*, 2023, 55(4): 188–196.]
- [10] Dong Xiangxiang, Battistelli G, Chisci L, et al. A variational Bayes moving horizon estimation adaptive filter with guaranteed stability[J]. *Automatica*, 2022, 142: 110374.
- [11] Narasimhappa M, Mahindrakar A D, Guizilini V C, et al. MEMS-based IMU drift minimization: Sage husa adaptive robust Kalman filtering[J]. *IEEE Sensors Journal*, 2020, 20(1): 250–260.
- [12] Tripathi R P, Singh A K, Gangwar P. Innovation-based fractional order adaptive Kalman filter[J]. *Journal of Electrical Engineering*, 2020, 71(1): 60–64.
- [13] Jiang Liuyang, Zhang Hai. Redundant measurement-based second order mutual difference adaptive Kalman filter[J]. *Automatica*, 2019, 100: 396–402.
- [14] Liu Tong, Zhang Zengjie, Liu Fangzhou, et al. Adaptive observer for a class of systems with switched unknown parameters using DREM[J]. *IEEE Transactions on Automatic Control*, 2024, 69(4): 2445–2452.
- [15] Zhang Xianku, Zhao Baigang, Zhang Guoqing. Improved parameter identification algorithm for ship model based on nonlinear innovation decorated by sigmoid function[J]. *Transportation Safety and Environment*, 2021, 3(2): 114–122.
- [16] Zhao Baigang, Zhang Xianku. An improved nonlinear innovation-based parameter identification algorithm for ship models[J]. *Journal of Navigation*, 2021, 74(3): 549–557.
- [17] Shi Zhenwei, Yang Haodong, Dai Mei. The data-filtering based bias compensation recursive least squares identification for multi-input single-output systems with colored noises[J]. *Journal of the Franklin Institute*, 2023, 360(7): 4753–4783.
- [18] Xu Huan, Ding Feng, Yang Erfu. Modeling a nonlinear process using the exponential autoregressive time series model[J]. *Nonlinear Dynamics*, 2019, 95(3): 2079–2092.
- [19] Ge Zhengwei, Ding Feng, Xu Ling, et al. Gradient-based iterative identification method for multivariate equation-error autoregressive moving average systems using the decomposition technique[J]. *Journal of the Franklin Institute*, 2019, 356(3): 1658–1676.
- [20] You Junyao, Liu Yanjun, Chen Jing, et al. Iterative identification for multiple-input systems with time-delays based on greedy pursuit and auxiliary model[J]. *Journal of the Franklin Institute*, 2019, 356(11): 5819–5833.
- [21] Lv Lei, Sun Wei, Pan Jian. Two-stage and three-stage recursive gradient identification of Hammerstein nonlinear systems based on the key term separation[J]. *International Journal of Robust and Nonlinear Control*, 2024, 34(2): 829–848.
- [22] Ding Feng, Xu Ling, Zhang Xiao, et al. Recursive identification methods for general stochastic systems with colored noises by using the hierarchical identification principle and the filtering identification idea[J]. *Annual Reviews in Control*, 2024, 57: 100942.
- [23] Yang Dan, Liu Yanjun, Ding Feng, et al. Hierarchical gradient-based iterative parameter estimation algorithms for a non-

- linear feedback system based on the hierarchical identification principle[J]. *Circuits, Systems, and Signal Processing*, 2024, 43(1):124–151.
- [24] Patil P V, Vachhani L, Ravitharan S, et al. Sequential state and unknown parameter estimation strategy and its application to a sensor fusion problem[J]. *IEEE Sensors Journal*, 2022, 22(21):20665–20675.
- [25] Molaei A, Nikoofard A, Sedigh A K, et al. Parameter and state estimation of managed pressure drilling system using the optimization-based supervisory framework[J]. *IEEE Transactions on Control Systems Technology*, 2023, 31(6): 2937–2944.
- [26] Aslan S, Cemgil A T, Akın A. Joint state and parameter estimation of the hemodynamic model by particle smoother expectation maximization method[J]. *Journal of Neural Engineering*, 2016, 13(4):046010.
- [27] Abolhasani M, Rahmani M. Robust deterministic least-squares filtering for uncertain time-varying nonlinear systems with unknown inputs[J]. *Systems & Control Letters*, 2018, 122:1–11.
- [28] Fernandes M R, do Val J B R, Souto R F. Robust estimation and filtering for poorly known models[J]. *IEEE Control Systems Letters*, 2020, 4(2):474–479.
- [29] Wang Ke, Wu Panlong, Li Xingxiu, et al. An adaptive outlier-robust Kalman filter based on sliding window and Pearson type VII distribution modeling[J]. *Signal Processing*, 2024, 216:109306.
- [30] Hua Jinxing, Liu Ruirui, Hao Fei. Two-channel false data injection attacks on multi-sensor remote state estimation[J]. *Asian Journal of Control*, 2023, 25(5):3776–3791.
- [31] Ge Quanbo, Ma Zhongcheng, Li Jinglan, et al. Adaptive cubature Kalman filter with the estimation of correlation between multiplicative noise and additive measurement noise[J]. *Chinese Journal of Aeronautics*, 2022, 35(5):40–52.
- [32] Xue Wei, Luan Xiaoli, Zhao Shunyi, et al. An online performance index for the Kalman filter[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71:1007912.
- [33] Zhang Tengfei, Jia Yingmin. Input-constrained optimal output synchronization of heterogeneous multiagent systems via observer-based model-free reinforcement learning[J]. *Asian Journal of Control*, 2024, 26(1):98–113.
- [34] Zhao Shunyi, Shmaliy Y S, Ahn C K, et al. Self-tuning unbiased finite impulse response filtering algorithm for processes with unknown measurement noise covariance[J]. *IEEE Transactions on Control Systems Technology*, 2021, 29(3):1372–1379.

Kalman State Estimation Algorithm Based on Q-learning Under Unknown Model Parameters

YANG Wenying, LUAN Xiaoli*, LIU Fei

(Key Laboratory for Advanced Process Control of Light Industry of Ministry of Education, Institute of Automation, Jiangnan University, Wuxi 214122, China)

Abstract:

Objective Kalman filtering (KF), as a widely used state estimation algorithm, plays a crucial role in estimating system state variables. The high-precision Kalman state estimation algorithm requires accurate knowledge of model parameters and noise statistical characteristics. Otherwise, estimation performance significantly degrades, and filter divergence can occur. However, in practical applications, many system model parameters and noise statistical characteristics are often unknown or inaccurate. Therefore, a Q-learning-based Kalman filtering (QL-KF) algorithm is proposed that learns and estimates simultaneously when model parameters and noise statistical characteristics are unknown.

Methods The Q-learning policy iteration algorithm, which was divided into two parts, policy improvement and policy evaluation, was employed to address the issue of unknown model information. In the policy improvement stage, a state-action value function (Q function) that evaluated the estimated state value was defined. Then, a formula transformation was utilized to ensure that the estimated value depended only on observed values rather than model parameters, eliminating the need for model parameters. In addition, two adjustable weight matrices were introduced to calculate the Kalman gain, avoiding reliance on the system noise statistical characteristics. Then, an estimation policy for obtaining system state estimates was derived from the Q function. In the policy evaluation stage, the estimation of the Q function was transformed into the estimation of its information matrix, and the recursive least squares algorithm was applied to identify the information matrix. Afterward, based on the identified information matrix, the estimation policy was followed to execute the corresponding actions and update the estimated values of the state variables. Finally, the proposed algorithm was applied to estimate the state of a two-state polynomial system and the water level of a quadruple water tank system to verify the effectiveness and feasibility of the algorithm. In addition, the proposed algorithm was compared to a joint state and parameter estimation algorithm.

Results and Discussions The estimation performance of the QL-KF algorithm was analyzed under conditions of unknown model parameters and noise statistical characteristics. A Monte Carlo experiment was conducted, and 50 Monte Carlo simulations were performed to enhance the credibility of the simulation. Uncertainty was introduced into each parameter to verify the robustness of the proposed algorithm. The root mean square

error (RMSE) and the average RMSE (ARMSE) were used as performance evaluation metrics. For the two-state polynomial system, when both the system process noise and measurement noise were Gaussian noise, the simulation results showed that the RMSE of the QL-KF algorithm exhibited a strong convergence trend, demonstrating the effectiveness of the algorithm. Because the initial estimates were randomly assigned and the Q-learning algorithm required some data accumulation during application, the initial RMSE was slightly larger and fluctuated, but showed a decreasing trend with an increasing number of iterations and gradually stabilized. Compared to the standard KF algorithm, when the model parameters were known in the KF algorithm, the RMSE value of the KF algorithm was low and very stable. However, when the model parameters of both algorithms were unknown, the proposed QL-KF algorithm achieved significantly better estimation accuracy than the standard KF algorithm, demonstrating stronger robustness. Compared to the EVIU algorithm (joint state and parameter estimation algorithm), the RMSE of the QL-KF algorithm was smaller and more stable after convergence, with an average ARMSE reduction of 34.66%, indicating higher estimation accuracy. It also demonstrated stronger robustness under parameter uncertainties. In addition, the algorithm required less computational time, reducing the average running time by 44.38%, and exhibited high real-time performance. When both the system process noise and measurement noise were non-Gaussian noise, the simulation results still showed that the RMSE of the QL-KF algorithm exhibited a convergence trend, confirming the algorithm's effectiveness. When the system parameters were unknown, the estimation error of the proposed algorithm was lower than that of the KF algorithm and similar to that of the EVIU algorithm. The running time of the QL-KF algorithm was reduced by 45.03% compared to the EVIU algorithm, indicating higher real-time performance. However, compared to the Gaussian noise system, the estimation error of the QL-KF system increased, indicating that different types of noise affected the estimation accuracy of the proposed algorithm. For the quadruple water tank system, the RMSE of the QL-KF algorithm for both state components showed favorable trends, demonstrating the effectiveness of the algorithm. Compared to the EVIU algorithm, the proposed algorithm exhibited stronger robustness under parameter uncertainties, with smaller estimation errors, an average ARMSE reduction of 79.93%, and a decrease in running time of 47.78%, indicating good real-time performance.

Conclusions The findings indicate that the proposed QL-KF algorithm can utilize only observations, without identifying system parameters, to estimate the internal state of systems when the model parameters and noise statistical characteristics are unknown. The estimation accuracy of the algorithm is influenced by the type of system noise. For Gaussian noise systems, the algorithm demonstrates high estimation accuracy, robust performance, and strong real-time capability. However, for non-Gaussian noise systems, the estimation accuracy decreases. Future work will focus on further improving estimation accuracy.

Key words: Kalman state estimation; Q-learning; unknown model parameters; unknown noise statistical characteristics

(编辑 张琼)

引用格式: Yang Wenying, Luan Xiaoli, Liu Fei. Kalman state estimation algorithm based on Q-learning under unknown model parameters[J]. *Advanced Engineering Sciences*, 2025, 57(6): 335-343. [杨雯莹, 栾小丽, 刘飞. 未知模型参数下基于Q学习的卡尔曼状态估计算法[J]. *工程科学与技术*, 2025, 57(6): 335-343.]