

不可达系统的鲁棒贝叶斯估计方法

易圣伦, 任雪梅[†]

(北京理工大学 自动化学院, 北京 100081)

摘要: 本文针对模型扰动下的不可达系统, 提出了一种新的针对退化分布下的极大极小博弈问题的求解和证明方法. 首先, 文章将有相对熵约束的极大极小博弈问题转换成了一个无约束的拉格朗日函数, 并找到其在均值和奇异的方差矩阵方向上都为严格凹函数的条件; 其次, 本文通过求解其均值和方差的极大值, 得到所对应的鲁棒贝叶斯估计器和奇异的扰动状态误差协方差矩阵; 最后, 文章证明存在一个唯一的拉格朗日乘子满足其约束条件. 微机电系统加速度计漂移估计仿真结果表明对所提算法的有效性.

关键词: 鲁棒估计; 贝叶斯理论; 不可达系统; 极大极小博弈

引用格式: 易圣伦, 任雪梅. 不可达系统的鲁棒贝叶斯估计方法. 控制理论与应用, 2024, 41(2): 221 – 228

DOI: 10.7641/CTA.2022.20177

Robust Bayesian estimation method for unreachable systems

YI Sheng-lun, REN Xue-mei[†]

(College of Automation, Beijing Institute of Technology, Beijing 100081, China)

Abstract: In this paper, a new solution and proof method are proposed for the minimax game problem under degenerate distribution. In particular, the minimax game problem subject to a relative entropy tolerance is first transformed into an unconstrained Lagrangian function. Accordingly, we tend to find the condition in which the unconstrained Lagrangian function is strictly concave along the direction of the singular variance matrix. Next, the robust Bayesian estimator and the disturbed state error covariance matrix are obtained by finding the maximizers of the corresponding mean and variance. Finally, it is shown that there is a unique Lagrange multiplier that satisfies our constraints. In addition, the proposed algorithm is applied to estimate the drift of MEMS accelerometers. The simulation results show that the proposed algorithm outperforms the standard Kalman filter.

Key words: robust estimation; Bayesian theory; unreachable system; minimax game

Citation: YI Shenglun, REN Xuemei. Robust Bayesian estimation method for unreachable systems. *Control Theory & Applications*, 2024, 41(2): 221 – 228

1 引言

鲁棒估计作为一种经典的估计方法, 被广泛地应用在控制^[1-2]、金融^[3-4]、通信^[5-6]、目标跟踪^[7-8]等领域. 其优点在于, 当待估系统模型出现扰动或者模型不匹配等情况时, 依然可以得到一个相对稳定且合理的估计值.

风险敏感滤波作为鲁棒估计中具有代表性的一类方法, 由 Speyer 等人^[9]在 1992 年的第 31 届决策与控制大会上提出. 此后, 由 Hansen 和 Levy 等人^[4, 10-11]逐渐完善并改进, 其核心思想在于采用了指数二次型函数的期望值作为目标函数, 替代了传统的标准二次型函数. 因此, 与贝叶斯估计器相比, 风险敏感滤波器在

Riccati 递推方程中增添了一个风险敏感因子, 可以对较大的误差加以更多的修正. 然而, 风险敏感滤波的缺点也是显而易见的, 该算法并没有给出风险敏感因子的解析解. 因此, 极大的依赖于扰动量的先验知识. 而在工业中, 诸如预知测量噪声分布或建模误差通常是不现实的. 因此, 2012 年, Levy 和 Nikoykhah^[12]提出了一种鲁棒状态空间滤波器, 该算法区别于风险敏感滤波算法中将模型扰动看作一个先验的参数, 转而用 Kullback-Leibler (KL) 散度来度量真实的模型和标称模型之间的不确定性. 由此, 可以通过给定一个比较好预知的模型扰动上界, 来求解所对应的风险敏感因子. 更重要的是, 本文可以通过构建一个以标称模

收稿日期: 2022-03-12; 录用日期: 2022-12-22.

[†]通信作者. E-mail: xmren@bit.edu.cn.

本文责任编辑: 张焕水.

国家自然科学基金项目(62273050, 61973036)资助.

Supported by the National Natural Science Foundation of China (62273050, 61973036).

型为球心, 模型扰动上界(也称容许误差)为半径的“球体”, 即包含所有可能的真实模型的模糊集, 来对真实模型进行约束, 从而找到其中最坏的模型(称作最不利模型), 然后再根据这个最不利模型来设计最优的估计器. 该估计器考虑了最坏的情况的最优估计器, 利用数学语言来说, 就是通过找出一个极大极小问题的鞍点, 从而得到所对应的最不利模型和最优鲁棒估计器. 此后, Zorzi^[13]设计出了基于Tau散度的约束下的鲁棒卡尔曼滤波器; 后来, 又将其思想扩展到分布式卡尔曼滤波上, 设计了鲁棒分布式卡尔曼滤波器^[14]. 同时, 国内也有学者对其展开的相关的研究^[15-16].

然而, 与标准的贝叶斯估计器或卡尔曼滤波器不同, 无论是风险敏感滤波器抑或是鲁棒状态空间滤波器, 都会在Riccati递推方程中涉及到求逆的过程, 而对于不可达的系统, 其状态向量 x 的分布为退化的, 这意味着, 无法利用KL散度去衡量模型间的不确定性, 同时也无法对状态误差协方差矩阵进行求逆操作. 因此针对以上问题, 本文基于文献[17-18]中所提出退化分布下的极大极小博弈问题, 设计了新的求解和证明方法, 解决了在实际应用场景中, 无法预先得知基变换中的变化矩阵的问题. 具体而言, 本文将带KL散度(也称为相对熵)约束的估计问题转换成了一个无约束的拉格朗日函数, 且通过对该函数的二阶变分找到使其在奇异的方差矩阵上为严格凹函数的条件. 值得注意的是, 这里的方差是指状态向量 x 和观测向量 y 的退化联合概率密度函数的方差矩阵, 所以其为奇异矩阵. 接下来, 本文通过求解拉格朗日函数的一阶变分等于0, 可以得到其均值和方差的最大值. 最后, 本文证明了存在一个唯一的拉格朗日乘子满足其约束条件. 最后, 本文在对微机电系统(micro-electro mechanical system, MEMS)中加速度计漂移校准的应用场景中验证了所提出的算法. 验证结果表明了本文设计的低秩鲁棒贝叶斯估计器可以针对不可达系统且在模型扰动的前提下, 可以得到比标准卡尔曼滤波更优的结果.

符号说明: $x \sim \mathcal{N}(m, K)$ 为 x 的标称高斯分布, 且其均值为 m , 协方差矩阵为 K . $x \sim \mathcal{N}(\tilde{m}, \tilde{K})$ 为 x 的真实高斯分布, 且其均值为 \tilde{m} , 协方差矩阵为 \tilde{K} . 则令

$$\mathcal{U} = \{m \in \mathbb{R}^{n+p}, \text{ s.t. } \tilde{m} - m \in \text{Im}(K)\},$$

$$\mathcal{V} = \{K \in \mathcal{Q}^{n+p}, \text{ s.t. } \text{Im}(K) = \text{Im}(\tilde{K})\},$$

其中: 矩阵的像、核和迹分别由 $\text{Im}(K)$, $\text{ker}(K)$ 和 $\text{tr}(K)$ 表示. 给定一个对称矩阵 K : $K > 0$ ($K \geq 0$)意味着 K 是正定(半正定)的; $\sigma_{\max}(K)$ 是 K 的最大特征值; 若 K 为奇异矩阵, 则 K^+ 和 $\det^+(K)$ 分别表示奇异矩阵 K 的伪逆和伪行列式, 其中映射 $K \mapsto K^+$ 的关于 K , $\delta K \in \mathcal{V}$ 的一阶变分为^[19]

$$-K^+(\delta K)K^+ + K^+K^+(\delta K)(I - KK^+) + (I - K^+K)(\delta K)K^+K^+,$$

而映射 $K \mapsto \det^+(K)$ 的关于 K , $\delta K \in \mathcal{V}$ 的一阶变分为

$$\det^+(K)\text{tr}(K^+\delta K),$$

其中: \mathcal{Q}^n 为维数为 n 的对称矩阵的向量空间; \mathcal{Q}_+^n 为 n 维正定对称矩阵的锥体, 且 $\bar{\mathcal{Q}}_+^n$ 表明其闭包. 主对角线上的元素是 a_1, a_2, \dots, a_n 的对角矩阵被定义为 $\text{diag}\{a_1, a_2, \dots, a_n\}$.

2 问题描述

考虑一个不可达的线性系统

$$\begin{cases} x_{t+1} = A_t x_t + B_t v_t, \\ y_t = C_t x_t + D_t v_t, \end{cases} \quad (1)$$

其中: $x_t \in \mathbb{R}^n$ 是状态向量; $y_t \in \mathbb{R}^p$ 是观测向量; $A_t \in \mathbb{R}^{n \times n}$, $C_t \in \mathbb{R}^{p \times n}$, $v_t \in \mathbb{R}^{n+p}$ 是标准的白噪声, 且包含了过程噪声和测量噪声. 此外, 由于系统不可达, 因此, 令 $r_t + p := \text{rank}([B_t^T \ D_t^T]^T) < n + p$, 则不失一般性的, 可以令 $[B_t^T \ D_t^T]^T \in \mathbb{R}^{(n+p) \times (r_t+p)}$ 是一个列满秩矩阵, 且 $B_t \in \mathbb{R}^{n \times (r_t+p)}$, $D_t \in \mathbb{R}^{p \times (r_t+p)}$. 同时, 令 $v_t \in \mathbb{R}^{r_t+p}$. 最后, 假设 $D_t D_t^T$ 是正定的矩阵.

注1 若系统中 \mathcal{Q}_+^n 内所有状态从原点到第 $t+1$ 步都可到达则称为该系统可达, 反之则不可达.

定义 $z_t := [x_{t+1}^T \ y_t^T]^T$, 令 $f_0(x_0)$ 为 x_0 的概率密度函数, $\tilde{\phi}_t(z_t|x_t)$ 为以 x_t 为条件的 z_t 的概率密度函数, 则模型(1)在有限区间 $0 \leq t \leq T$ 上可以等效表示为一个标称联合概率密度函数^[17-18], 即

$$f(X_{T+1}, Y_T) = f_0(x_0) \prod_{t=0}^T \phi_t(z_t|x_t), \quad (2)$$

其中:

$$X_{T+1} = [x_0^T \ \dots \ x_t^T \ \dots \ x_{T+1}^T],$$

$$Y_T = [y_0^T \ \dots \ y_t^T \ \dots \ y_T^T],$$

值得注意的是, 因为 $r_t < n$, 所以 $\phi_t(z_t|x_t)$ 是一个退化的条件概率密度函数, 则令 \mathcal{A}_t 表示 ϕ_t 的支撑集, 且 \mathcal{A}_t 和 x_t 相关. 相应地, 假设真实模型在有限区间 $0 \leq t \leq T$ 上可以等效于如下真实的联合概率密度函数来表示:

$$\tilde{f}(X_{T+1}, Y_T) = \tilde{f}_0(x_0) \prod_{t=0}^T \tilde{\phi}_t(z_t|x_t). \quad (3)$$

此外, 假设 $\tilde{\phi}_t(z_t|x_t)$ 的支撑集为 \mathcal{A}_t , 则 $f(X_{T+1}, Y_T)$ 和 $\tilde{f}(X_{T+1}, Y_T)$ 就有了相同的支撑集, 称作 \mathcal{A} . 值得注意的是, KL散度是无法描述确定性模型的“距离”的, 因此, 引入文献[12, 17-18]的思想, 可以通过考虑以 \mathcal{A} 为支撑集的KL散度来描述真实模型和标称模型之间的模型不匹配程度, 即

$$D(\tilde{f}, f) = \int_{\mathcal{A}} \tilde{f} \ln \frac{\tilde{f}}{f} dX_{T+1} dY_T. \quad (4)$$

假定标称模型(1)已知, 则已知 $f(X_{T+1}, Y_T)$, 但是真实模型是未知的. 给定的观测数据 Y_{t-1} 可知, 可以假设 $\tilde{\phi}_t$ 属于接下来的凸的模糊集内^[17-18], 即

$$\mathcal{B}_t := \{ \tilde{\phi}_t, \text{ s.t. } \tilde{E}[\ln \frac{\tilde{\phi}_t(z_t|x_t)}{\phi_t(z_t|x_t)} | Y_{t-1}] \leq c_t \}, \quad (5)$$

其中

$$\begin{aligned} \tilde{E}[\ln \frac{\tilde{\phi}_t(z_t|x_t)}{\phi_t(z_t|x_t)} | Y_{t-1}] := \\ \int_{\tilde{\mathcal{A}}_t} \int_{\mathcal{A}_t} \tilde{\phi}_t(z_t|x_t) \tilde{f}_t(x_t|Y_{t-1}) \times \\ \ln \frac{\tilde{\phi}_t(z_t|x_t)}{\phi_t(z_t|x_t)} dz_t dx_t, \end{aligned} \quad (6)$$

且 $\tilde{\mathcal{A}}_t$ 表示 $\tilde{f}_t(x_t|Y_{t-1})$ 的支撑集, c_t 为容许误差, 即模型不确定程度的上限. 通俗的解释, 就是利用在KL散度的拓扑学原理, 形成一个以标称模型为球心, c_t 为半径的“球”, 如图1所示. 而值得注意的是, 由于模型(1)为一阶马尔可夫模型, 所以 $\tilde{\phi}_t(z_t|x_t)$ 和 $\tilde{\phi}_t(z_t|x_t, Y_{t-1})$ 是等效的.

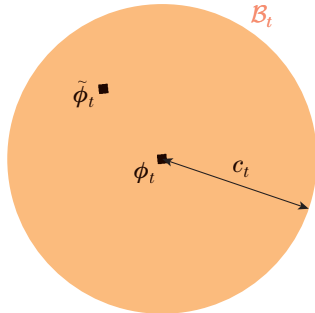


图1 无限时间的KL散度约束

Fig. 1 KL divergence constraint over infinite horizon

考虑设计在最坏的情况下的最优的估计器, 即在给定 Y_t 和 ϕ_t 的情况下, 求解在模糊集 \mathcal{B}_t 内 $\tilde{\phi}_t$ 的最大值以及相应的 x_{t+1} 的鲁棒最优估计器. 因此, 寻求解决如下动态极大极小问题^[17-18]:

$$\hat{x}_{t+1} = \arg \min_{g_t \in \mathcal{G}_t} \max_{\tilde{\phi}_t \in \mathcal{B}_t} J_t(\tilde{\phi}_t, g_t), \quad (7)$$

其中

$$\begin{aligned} J_t(\tilde{\phi}_t, g_t) = \frac{1}{2} \tilde{E}[\|x_{t+1} - g_t(y_t)\|^2 | Y_{t-1}] = \\ \frac{1}{2} \int_{\tilde{\mathcal{A}}_t} \int_{\mathcal{A}_t} \|x_{t+1} - g_t(y_t)\|^2 \times \\ \tilde{\phi}_t(z_t|x_t) \tilde{f}_t(x_t|Y_{t-1}) dz_t dx_t. \end{aligned} \quad (8)$$

\mathcal{G}_t 表示相对于所有密度 $\tilde{\phi}_t(z_t|x_t) \tilde{f}_t(x_t|Y_{t-1})$ 使得 $\tilde{\phi}_t \in \mathcal{B}_t$ 具有有限二阶矩的估计量集. 最后注意 $\tilde{\phi}_t$ 必须满足如下的约束:

$$I_t(\tilde{\phi}_t) \triangleq \int_{\tilde{\mathcal{A}}_t} \int_{\mathcal{A}_t} \tilde{\phi}_t(z_t|x_t) \tilde{f}_t(x_t|Y_{t-1}) dz_t dx_t = 1. \quad (9)$$

3 低秩鲁棒贝叶斯估计器

这一节主要研究通过求解极大极小问题(7)来设计鲁棒贝叶斯估计器. 不难看出, 其中 J 的鞍点必须满足如下条件.

对于所有的 $\tilde{f} \in \mathcal{B}_t$ 和 $g \in \mathcal{G}_t$,

$$J(\tilde{\phi}_t, g_t^0) \leq J(\tilde{\phi}_t^0, g_t^0) \leq J(\tilde{\phi}_t^0, g_t) \quad (10)$$

都成立. 首先, 找到式(8)中第1个不等式所成立的条件. 这里, 由文献[17]可知如下引理1.

引理1^[17] 对于一个固定的估计器 $g_t \in \mathcal{G}_t$, 对应的其支撑集为 \mathcal{A}_t 的真实退化概率密度函数 $\tilde{\phi}_t(z_t|x_t) \in \mathcal{B}_t$ 可使在约束条件 $D(\tilde{\phi}_t, \phi_t) \leq c_t$ 下的目标函数

$$J_t(\tilde{\phi}_t, g_t) = \tilde{E}[\|x_{t+1} - g_t(y_t)\|^2 | Y_{t-1}], \quad (11)$$

最大化的极值为

$$\tilde{\phi}_t^0 = \frac{1}{M_t(\lambda_t)} \exp\left(\frac{1}{2\lambda_t} \|x_{t+1} - g_t(y_t)\|^2\right) \phi_t, \quad (12)$$

其中 $\tilde{\phi}_t^0$ 为最不利概率密度函数. 此外, $M_t(\lambda_t)$ 是使式(9)成立的归一化常数. 最后, 如果 $c_t > 0$ 足够小, 将存在一个唯一的 $\lambda_t > 0$ 使得 $D(\tilde{\phi}_t^0, \phi_t) = c_t$.

其次, 需要找到式(8)中第2个不等式所成立的条件. 基于引理1可以得到最不利概率密度函数 $\tilde{\phi}_t^0$, 则其使目标函数 $J_t(\tilde{\phi}_t^0, g_t)$ 最小化的对应的贝叶斯估计器 $g_t \in \mathcal{G}_t$ 为^[12, 17-18]

$$\begin{aligned} \hat{x}_{t+1} = g_t^0(y_t) = \tilde{E}[x_{t+1} | Y_t] = \\ \int_{\mathcal{A}_t^*} x_{t+1} \tilde{f}_{t+1}(x_{t+1} | Y_t) dx_{t+1}, \end{aligned} \quad (13)$$

其中

$$\begin{aligned} \tilde{f}_{t+1}(x_{t+1} | Y_t) = \\ \frac{\int_{\tilde{\mathcal{A}}_t} \tilde{\phi}_t^0(z_t|x_t) \tilde{f}_t(x_t|Y_{t-1}) dx_t}{\int_{\tilde{\mathcal{A}}_t} \int_{\mathcal{A}_t^*} \tilde{\phi}_t^0(z_t|x_t) \tilde{f}_t(x_t|Y_{t-1}) dx_{t+1} dx_t}, \end{aligned}$$

\mathcal{A}_t^* 是 $\tilde{f}_t(x_{t+1}|Y_t)$ 的支撑集且 \mathcal{A}_t^{**} 有如下定义: 对于 $[x_{t+1}^T \ y_t^T]^T \in \mathcal{A}_t$, 当且仅当存在至少一个 y_t , 则 $x_{t+1} \in \mathcal{A}_t^{**}$.

由式(13)可以很明显地看出, 贝叶斯估计器 \hat{x}_{t+1} 是和最不利概率密度函数 $\tilde{\phi}_t^0(z_t|x_t)$ 相关的. 同时, 从式(12)也可以明显地看出, 最不利概率密度函数 $\tilde{\phi}_t^0(z_t|x_t)$ 同样取决于贝叶斯估计器 \hat{x}_{t+1} . 因此, 为了打破这个死循环, 需要加入一个假设条件: 假设在 t 时刻, 给定 Y_{t-1} , x_t 的先验条件密度为

$$\tilde{f}_t(x_t|Y_{t-1}) \sim \mathcal{N}(\hat{x}_t, \tilde{P}_t), \quad (14)$$

且其秩为 $\text{rank}(\tilde{P}_t) = r_t$. 此外, 引入不可达系统(1), 可以得到其标称的条件转移概率密度函数为

$$\phi_t(z_t|x_t) \sim \mathcal{N}(m_{z_t|x_t}, K_{z_t|x_t}), \quad (15)$$

且

$$m_{z_t|x_t} = \begin{bmatrix} A_t \\ C_t \end{bmatrix} x_t, \quad K_{z_t|x_t} = \begin{bmatrix} B_t B_t^T & B_t D_t^T \\ D_t B_t^T & D_t D_t^T \end{bmatrix}. \quad (16)$$

正如已经注意到的, 因为系统(1)不可达, 则后者是退化的. 因此, $K_{z_t|x_t}$ 为奇异矩阵. 这里, 定义标称边缘概率密度函数为

$$f_t(z_t|Y_{t-1}) = \int_{\tilde{\mathcal{A}}_t} \phi_t(z_t|x_t) \tilde{f}_t(x_t|Y_{t-1}) dx_t, \quad (17)$$

由于 $\phi_t(z_t|x_t)$ 和 $\tilde{f}_t(x_t|Y_{t-1})$ 都为高斯函数, 因此, 可以得知 $f_t(z_t|Y_{t-1}) \sim \mathcal{N}(m_{z_t}, K_{z_t})$, 且

$$m_{z_t} = \begin{bmatrix} A_t \\ C_t \end{bmatrix} \hat{x}_t, \quad K_{z_t} = \begin{bmatrix} K_{x_{t+1}} & K_{x_{t+1}y_t} \\ K_{y_t x_{t+1}} & K_{y_t} \end{bmatrix}, \quad (18)$$

其中 K_{z_t} 的参数形式如下:

$$K_{z_t} = \begin{bmatrix} A_t \\ C_t \end{bmatrix} \tilde{P}_t [A_t^T \quad C_t^T] + \begin{bmatrix} B_t \\ D_t \end{bmatrix} [B_t^T \quad D_t^T]. \quad (19)$$

则有

$$K_{y_t} = C_t \tilde{P}_t C_t^T + D_t D_t^T \geq D_t D_t^T > 0. \quad (20)$$

接下来, 定义真实概率密度函数为

$$\tilde{f}_t(z_t|Y_{t-1}) = \int_{\tilde{\mathcal{A}}_t} \phi_t(z_t|x_t) \tilde{f}_t(x_t|Y_{t-1}) dx_t, \quad (21)$$

其中: 真实概率密度函数和 $f_t(z_t|Y_{t-1})$ 有相同的支持集, 且将其定义为: $\tilde{\mathcal{A}}_t$. 则不难看出, 如果 $\phi_t(z_t|x_t)$ 是高斯函数, 则标称边缘分布对应的最不利边缘概率密度函数

$$\tilde{f}_t^0(z_t|Y_{t-1}) = \frac{1}{M(\lambda_t)} \exp\left(\frac{1}{2\lambda_t} \|x_{t+1} - g_t(y_t)\|^2\right) \times f_t(z_t|Y_{t-1}) \quad (22)$$

也为高斯. 因为 $f_t(z_t|Y_{t-1})$ 和 $\tilde{f}_t^0(z_t|Y_{t-1})$ 为高斯, 则有如下引理2.

引理 2 给定一个观测向量 $y \in \mathbb{R}^p$, 随机待估向量 $x \in \mathbb{R}^n$. 假设 x 和 y 的联合标称分布 $f(z)$ 是高斯分布的, 即

$$z = \begin{bmatrix} x \\ y \end{bmatrix} \sim \mathcal{N}(m_z, K_z), \quad (23)$$

其中 $m_z = \begin{bmatrix} m_x \\ m_y \end{bmatrix}$ 和 $K_z = \begin{bmatrix} K_x & K_{xy} \\ K_{yx} & K_y \end{bmatrix}$ 分别表示 z_t 的均值向量和协方差矩阵. 则贝叶斯估计器为

$$\hat{x} = m_x + K_{xy} K_y^{-1} (y - m_y), \quad (24)$$

因此, 可以得到鲁棒贝叶斯估计器和奇异的扰动状态误差协方差矩阵, 如定理1所示.

定理 1 最不利高斯概率密度函数 $\tilde{f}_t^0(z_t|Y_{t-1})$

有如下的均值和方差:

$$\tilde{m}_{z_t}^0 = m_{z_t} = \begin{bmatrix} m_{x_t} \\ m_{y_t} \end{bmatrix}, \quad \tilde{K}_{z_t}^0 = \begin{bmatrix} \tilde{K}_{x_t} & K_{x_t y_t} \\ K_{y_t x_t} & K_{y_t} \end{bmatrix}, \quad (25)$$

不难看出, 有且仅有 x_t 是扰动的.

则基于引理2, 可设计对应的贝叶斯估计器

$$g_t^0(y_t) = G_t(y_t - m_{y_t}) + m_{x_t}, \quad (26)$$

且 $G_t = K_{x_t y_t} K_{y_t}^{-1}$.

此外, 给定 Y_{t-1}, x_t 的标称后验协方差矩阵(即 Riccati 方程)为

$$P_t := K_{x_t} - K_{x_t y_t} K_{y_t}^{-1} K_{y_t x_t}, \quad (27)$$

其最不利的后验协方差矩阵为

$$\tilde{P}_t = (P_t^+ - \lambda^{-1} H_t^T H_t)^+, \quad (28)$$

其中 $H_t^T \in \mathbb{R}^{n \times r_t}$ 是一个矩阵, 其列对于 $\text{Im}(P_t)$ 形成一个正交基. 此外, 还存在一个唯一的拉格朗日乘子 $\lambda > \sigma_{\max}(\Lambda_{P_t}) > 0$ 使得 $c_t = D(\tilde{f}_t^0, f_t)$, 其中

$$\Lambda_{P_t} := H_t H_t^T (H_t P_t^+ H_t^T)^{-1} H_t H_t^T. \quad (29)$$

证 J 的鞍点必须满足如下条件: 对于所有的 $\tilde{f} \in \mathcal{B}$ 和 $g \in \mathcal{G}$, 找到 $J(\tilde{f}^0, g^0)$ 使

$$J(\tilde{f}, g^0) \leq J(\tilde{f}^0, g^0) \leq J(\tilde{f}^0, g), \quad (30)$$

首先, 证明式(30)中第1个不等式.

步骤 1 首先, \tilde{f} 由均值向量 \tilde{m}_z 和协方差矩阵 \tilde{K}_z 两个变量组成, 则给定一个固定的 g^0 , 有

$$J(\tilde{f}, g^0) = \frac{1}{2} \text{tr} \left(\begin{bmatrix} I_n \\ -G_0^T \end{bmatrix} H^T H [I_n - G_0] \times (\tilde{K}_z + \Delta m_z \Delta m_z^T) \right). \quad (31)$$

不难看出, J 的参数化形式由 \tilde{m}_z 和 \tilde{K}_z 构成, 同时, 将 f 和 \tilde{f} 的退化高斯函数带入式(4)中, 易知 KL 散度的参数化解析表达式为

$$D(\tilde{f}, f) = \frac{1}{2} (\Delta m_z^T K_z^+ \Delta m_z + \ln \det^+(K_z) - \ln \det^+(\tilde{K}_z) + \text{tr}(K_z^+ \tilde{K}_z) - (r + p)), \quad (32)$$

其中 $\Delta m_z = \tilde{m}_z - m_z$. 因此, 关于 $\tilde{f} \in \mathcal{B}$, 使 $J(\tilde{f}, g^0)$ 最大化的求解可以等效于在 $D(\tilde{f}, f) \leq c$ 的约束下, 关于 $\tilde{m}_z \in \mathcal{U}$ 和 $\tilde{K}_z \in \mathcal{V}$ 使 $J(\tilde{f}, g^0)$ 最大化求解.

因此, 可以构造其相关的拉格朗日函数为

$$L(\tilde{m}_z, \tilde{K}_z, \lambda) = J(\tilde{f}, g^0) + \lambda(c - D(\tilde{f}, f)) = \frac{1}{2} \text{tr} \left\{ \left(\begin{bmatrix} I_n \\ -G_0^T \end{bmatrix} H^T H [I_n - G_0] - \lambda K_z^+ \right) \tilde{K}_z \right\} + \frac{1}{2} \Delta m_z^T \left(\begin{bmatrix} I_n \\ -G_0^T \end{bmatrix} H^T H [I_n - G_0] - \lambda K_z^+ \right) \times$$

$$\Delta m_z - \frac{\lambda}{2} \ln \det^+(K_z) + \frac{\lambda}{2} \ln \det^+(\tilde{K}_z) + \lambda c + \frac{\lambda}{2}(r+p), \quad (33)$$

其中 $\lambda \geq 0$ 为和约束相关的拉格朗日乘子。

其次, 考虑关于 $\tilde{m}_z \in \mathcal{U}$ 和 $\tilde{K}_z \in \mathcal{V}$ 使 $L(\tilde{m}_z, \tilde{K}_z, \lambda)$ 的无约束最大化求解, 其中:

$$\mathcal{U} = \{\tilde{m}_z \in \mathbb{R}^{n+p}, \text{ s.t. } \tilde{m}_z - m_z \in \text{Im}(K_z)\}, \quad (34)$$

$$\mathcal{V} = \{\tilde{K}_z \in \mathcal{Q}^{n+p}, \text{ s.t. } \text{Im}(K_z) = \text{Im}(\tilde{K}_z)\}. \quad (35)$$

则关于 $\tilde{m}_z \in \mathcal{U}$, 令

$$W(\lambda) \triangleq \begin{bmatrix} I_n \\ -G_0^T \end{bmatrix} H^T H [I_n - G_0] - \lambda K_z^+. \quad (36)$$

因此, 在 $\tilde{m}_z \in \mathcal{U}$ 方向上的 $L(\tilde{m}_z, \tilde{K}_z, \lambda)$ 的一阶变分和二阶变分分别如下:

$$\delta L(\tilde{m}_z, \tilde{K}_z, \lambda; \delta \tilde{m}_z) = \frac{1}{2} \delta \Delta m_z^T W(\lambda) \Delta m_z + \frac{1}{2} \Delta m_z^T W(\lambda) \delta \Delta m_z, \quad (37)$$

$$\delta^2 L(\tilde{m}_z, \tilde{K}_z, \lambda; \delta \tilde{m}_z, \delta \tilde{m}_z) = \delta \tilde{m}_z^T W(\lambda) \delta \tilde{m}_z, \quad (38)$$

因此, 当且仅当 $W(\lambda)$ 是半负定矩阵, 且 $\text{Im}(W(\lambda)) = \text{Im}(K_z)$ 时, $L(\tilde{m}_z, \tilde{K}_z, \lambda)$ 是严格的凹函数。

其次, 需要找到使 $W(\lambda)$ 为负半定矩阵的 λ 的条件。值得注意的是, K_z 可以被分解为一个上三角矩阵、一个斜对角矩阵和一个下三角矩阵(UDL(universal data link)分解), 即

$$K_z = \begin{bmatrix} I_n & G_0 \\ 0 & I_p \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & K_y \end{bmatrix} \begin{bmatrix} I_n & 0 \\ G_0^T & I_p \end{bmatrix}, \quad (39)$$

由于 $\text{rank}(K_z) = r + p$ 且 $K_y > 0$, 鉴于式(39)可知: $\text{rank}(P) = r$. 因此, 它的伪逆可以进行以下UDL分解:

$$K_z^+ = \left(\begin{bmatrix} I_n & G_0 \\ 0 & I_p \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & K_y \end{bmatrix} \begin{bmatrix} I_n & 0 \\ G_0^T & I_p \end{bmatrix} \right)^+ = \begin{bmatrix} I_n & 0 \\ -G_0^T & I_p \end{bmatrix} \begin{bmatrix} P^+ & 0 \\ 0 & K_y^{-1} \end{bmatrix} \begin{bmatrix} I_n & -G_0 \\ 0 & I_p \end{bmatrix}. \quad (40)$$

因此, 不难看出

$$W(\lambda) = \begin{bmatrix} I_n & 0 \\ -G_0^T & I_p \end{bmatrix} \begin{bmatrix} H^T H - \lambda P^+ & 0 \\ 0 & -\lambda K_y^{-1} \end{bmatrix} \begin{bmatrix} I_n & -G_0 \\ 0 & I_p \end{bmatrix}. \quad (41)$$

则 $W(\lambda) \leq 0$, 且 $\text{Im}(W(\lambda)) = \text{Im}(K_z)$ 当且仅当

$$H^T H - \lambda P^+ \leq 0, \text{ Im}(H^T H - \lambda P^+) = \text{Im}(P^+). \quad (42)$$

由于 $\text{Im}(H^T) \subseteq \text{Im}(P) = \text{Im}(P^+)$, 上述条件可以等

效于条件

$$H(H^T H - \lambda P^+) H^T < 0, \quad (43)$$

这意味着

$$I_q - \lambda(HH^T)^{-1} H P^+ H^T (HH^T)^{-1} < 0, \quad (44)$$

因此, $W(\lambda)$ 为半负定矩阵当且仅当

$$\lambda > \sigma_{\max}(\Lambda_P). \quad (45)$$

其次, 求得 $L(\tilde{m}_z, \tilde{K}_z, \lambda)$ 在 $\delta \tilde{K}_z \in \mathcal{V}$ 方向上的一阶变分和二阶变分分别如下:

$$\begin{aligned} \delta L(\tilde{m}_z, \tilde{K}_z, \lambda; \delta \tilde{K}_z) &= \frac{1}{2} \text{tr}(W(\lambda) \delta \tilde{K}_z) + \frac{\lambda}{2} \frac{\partial \ln \det^+(\tilde{K}_z)}{\partial \det^+(\tilde{K}_z)} \delta \det^+(\tilde{K}_z; \delta \tilde{K}_z) = \frac{1}{2} \text{tr}(W(\lambda) \delta \tilde{K}_z) + \frac{\lambda}{2} \frac{1}{\det^+(\tilde{K}_z)} \det^+(\tilde{K}_z) \text{tr}(\tilde{K}_z^+ \delta \tilde{K}_z) = \frac{1}{2} \text{tr}(W(\lambda) + \lambda \tilde{K}_z^+) \delta \tilde{K}_z, \end{aligned} \quad (46)$$

且

$$\begin{aligned} \delta^2 L(\tilde{m}_z, \tilde{K}_z, \lambda; \delta \tilde{K}_z, \delta \tilde{K}_z) &= -\frac{\lambda}{2} \text{tr}(\delta \tilde{K}_z \tilde{K}_z^+ \delta \tilde{K}_z \tilde{K}_z^+) \leq 0, \end{aligned}$$

因此, $L(\tilde{m}_z, \tilde{K}_z, \lambda)$ 是在 $\tilde{K}_z \in \mathcal{V}$ 方向上的严格凹函数。

步骤 2 求解其极大值点. 对于均值而言, 当且

仅当

$$\Delta m_z = 0, \quad (47)$$

关于 $\tilde{m}_z \in \mathcal{U}$, $L(\tilde{m}_z, \tilde{K}_z, \lambda)$ 取得极大值, 即, $\tilde{m}_z = m_z$.

对于协方差矩阵而言, 则可以通过条件 $\delta L(\tilde{m}_z, \tilde{K}_z, \lambda; \delta \tilde{K}_z) = 0, \forall \delta \tilde{K}_z \in \mathcal{V}$ 得到. 基于 $\lambda > 0$ 的假设, 其极大值为

$$\begin{aligned} \tilde{K}_z^+ &= -\frac{1}{\lambda} W(\lambda) = K_z^+ - \frac{1}{\lambda} \begin{bmatrix} I_n \\ -G_0^T \end{bmatrix} H^T H [I_n \ G_0]. \end{aligned} \quad (48)$$

接下来, 取代式(40)中的式(48)有

$$\begin{aligned} \tilde{K}_z^+ &= \begin{bmatrix} I_n & 0 \\ -G_0^T & I_p \end{bmatrix} \begin{bmatrix} P^+ & 0 \\ 0 & K_y^{-1} \end{bmatrix} \begin{bmatrix} I_n & -G_0 \\ 0 & I_p \end{bmatrix} - \frac{1}{\lambda} \begin{bmatrix} I_n \\ -G_0^T \end{bmatrix} H^T H [I_n \ G_0] = \end{aligned}$$

$$\begin{bmatrix} I_n & 0 \\ -G_0^T & I_p \end{bmatrix} \begin{bmatrix} P^+ - \lambda^{-1}H^T H & 0 \\ 0 & K_y^{-1} \end{bmatrix} \begin{bmatrix} I_n & -G_0 \\ 0 & I_p \end{bmatrix}, \quad (49)$$

且在条件(45)下是半正定的. 由此可知

$$\tilde{K}_z = \begin{bmatrix} I_n & G_0 \\ 0 & I_p \end{bmatrix} \begin{bmatrix} \tilde{P} & 0 \\ 0 & K_y \end{bmatrix} \begin{bmatrix} I_n & 0 \\ G_0^T & I_p \end{bmatrix}, \quad (50)$$

其中

$$\tilde{P} = (P^+ - \lambda^{-1}H^T H)^+, \quad (51)$$

且 $\text{rank}(\tilde{P}) = r$.

步骤 3 最后, 还需要证明存在一个唯一的拉格朗日乘子 $\lambda > \sigma_{\max}(A_P) > 0$, 满足条件 $D(\tilde{f}^0, f) = c$. 令

$$P = U M U^T \quad (52)$$

为 P 的特征值分解, 其中 $U \in \mathbb{R}^{n \times n}$ 使得 $U^T U = U U^T = I_n$ 和

$$M = \text{diag}\{m_1, \dots, m_r, 0, \dots, 0\}$$

且 $m_i > 0$. 因此, 有 $P^+ = U M^+ U^T$. 由于 $\text{Im}(H^T) \subseteq \text{Im}(P)$, 可知

$$H^T H = U D U^T, \quad (53)$$

其中 $D \in \bar{\mathcal{Q}}_+^n$, 且不失一般性的, 它可以被分割为

$$D = \begin{bmatrix} D_* & 0 \\ 0 & 0 \end{bmatrix},$$

且 $D_* \in \bar{\mathcal{Q}}_+^{r \times r}$; $\text{rank}(D_*) = q$. 由式(51)–(53)可知

$$\tilde{P} = U(M^+ - \lambda^{-1}D)^+ U^T, \quad (54)$$

同时将式(47)(40)(50)(54)代入式(32)中, 则不难得出

$$\begin{aligned} D(\tilde{f}^0, f) &= \\ \frac{1}{2} \{ \ln \det^+(P) - \ln \det^+(\tilde{P}) + \text{tr}(P^+ \tilde{P} - I_r) \} &= \\ \frac{1}{2} \{ \ln \det^+(M) + \ln \det^+(M^+ - \lambda^{-1}D) + & \\ \text{tr}[M^+(M^+ - \lambda^{-1}D)^+ - I_r] \} &= \\ \frac{1}{2} \{ \ln \det(M_*) + \ln \det(M_*^{-1} - \lambda^{-1}D_*) + & \\ \text{tr}[M_*^{-1}(M_*^{-1} - \lambda^{-1}D_*)^{-1} - I_r] \}, & \quad (55) \end{aligned}$$

其中 $M_* = \text{diag}\{m_1, \dots, m_r\}$. 这里, 值得注意的是: $M_*^{-1} - \lambda^{-1}D_*$ 在条件式(45)下为一个正定矩阵. 接下来, 用符号 $\mu(\lambda) := D(\tilde{f}^0, f)$ 来表示 $D(\tilde{f}^0, f)$ 对 λ 的相关性. 则关于 λ , $\mu(\lambda)$ 的一阶变分为

$$\begin{aligned} \delta \mu(\lambda; \delta \lambda) &= \frac{1}{2} \{ \delta \lambda \text{tr}((M_*^{-1} - \lambda^{-1}D_*)^{-1} \lambda^{-2} D_*) - \\ \delta \lambda \text{tr}(M_*^{-1} (M_*^{-1} - \lambda^{-1}D_*)^{-1} \times & \\ \lambda^{-2} D_* (M_*^{-1} - \lambda^{-1}D_*)^{-1}) \}. & \quad (56) \end{aligned}$$

因此

$$\begin{aligned} \frac{\delta \mu(\lambda; \delta \lambda)}{\delta \lambda} &= \\ \frac{\lambda^{-2}}{2} \{ \text{tr}((M_*^{-1} - \lambda^{-1}D_*)^{-1} D_* - & \\ (M_*^{-1} - \lambda^{-1}D_*)^{-1} D_* \times & \\ (M_*^{-1} - \lambda^{-1}D_*)^{-1} M_*^{-1}) \} &= \\ \frac{\lambda^{-2}}{2} \{ \text{tr}((M_*^{-1} - \lambda^{-1}D_*)^{-1} D_* \times & \\ (I_r - (M_*^{-1} - \lambda^{-1}D_*)^{-1} M_*^{-1})) \} &= \\ \frac{\lambda^{-2}}{2} \{ \text{tr}((M_*^{-1} - \lambda^{-1}D_*)^{-1} D_* \times & \\ (M_* (M_*^{-1} - \lambda^{-1}D_*) - I_r) \times & \\ (M_*^{-1} - \lambda^{-1}D_*)^{-1} M_*^{-1}) \} &= \\ -\frac{\lambda^{-3}}{2} \{ \text{tr}(M_*^{-\frac{1}{2}} (M_*^{-1} - \lambda^{-1}D_*)^{-1} \times & \\ D_* M_* D_* (M_*^{-1} - \lambda^{-1}D_*)^{-1} M_*^{-\frac{1}{2}}) \} &< 0. \end{aligned}$$

因为 M_* , $M_*^{-1} - \lambda^{-1}D_*$ 为正定矩阵, 由此, 迹内的矩阵是正半定的, 且与零矩阵不同. 因此, $\mu(\lambda)$ 是严格单调递减的. 此外, 不难看出

$$\lim_{\lambda \rightarrow \infty} \mu(\lambda) = 0, \quad \lim_{\lambda \downarrow \sigma_{\max}(A_P)} \mu(\lambda) = \infty. \quad (57)$$

由此可得到结论: 存在一个唯一的拉格朗日乘子 $\lambda > \sigma_{\max}(A_P) > 0$, 使得 $D(\tilde{f}^0, f) = c$.

最后, 式(24)中右边的不等式可以基于引理2得到, 所求得的鲁棒贝叶斯估计器由式(20)所示. 证毕.

注 2 算法稳定性和收敛性分析可参考文献[17].

4 仿真与分析

市面上常用的大部分智能移动设备同时集成了MEMS微机电系统芯片和GNSS信号接收芯片^[20]. 但是, MEMS加速度计的漂移会导致很大的误差, 因此, MEMS加速度计漂移校准一直是一个热门的话题. 这里, 本文考虑基于GPS测量数据的MEMS加速度漂移估计模型^[21–22]

$$\begin{cases} x_{t+1} = A x_t + U u_t + B \omega_t, \\ y_t = C x_t + D \omega_t, \end{cases} \quad (58)$$

其中: 状态向量为 $x_t = [s_t \ v_t \ d]^T$; s_t 和 v_t 分别表示目标的位移和速度. 相应地, d 是MEMS加速度计所测得的加速度漂移, 这里, 将其看作一个常数. 此外, 输入为MEMS加速度计所测得的加速度 $u_t = a_t$. 接下来, 测量向量为 $y_t = [s_t^G \ v_t^G]^T$, 且 s_t^G 和 v_t^G 分别为GPS所测的位置信息和速度信息. 其次, ω_t 为一个4维的高斯标准白噪声. 此外, 过程矩阵(包含状态转移矩阵 A 和输入矩阵 U)为

$$A = \begin{bmatrix} 1 & T & -0.5T^2 \\ 0 & 1 & -T \\ 0 & 0 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 0.5T^2 \\ T \\ 0 \end{bmatrix}, \quad (59)$$

其中 $T = 0.1$ 为采样周期; 而测量矩阵为

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \quad (60)$$

最后, 过程噪声矩阵和测量噪声矩阵分别为

$$B = \begin{bmatrix} 0.2 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (61)$$

值得注意的是, 这个模型是不可达的, 即 $\text{rank}([B^T D^T]^T) = 4) < n + p$. 因此, 传统的鲁棒滤波算法, 如风险敏感滤波^[20], 鲁棒最小二乘^[10], 鲁棒状态空间滤波^[12]等, 将不在适用于这个应用场景. 这里, 为了验证所提算法的有效性, 需要设计上述标称模型所对应的含有扰动的真实模型. 然而, 对于传感器漂移而言, 其扰动很难用有限个参数来描述, 因此, 基于本文利用相对熵函数来描述该非参数化扰动. 具体来说, 引入文献[17]第5章中的最不利模型, 即给定一个容许误差 c , 所有可能的真实模型中最坏的一个. 具体如下.

首先, 取状态空间向量

$$\xi_t \triangleq \begin{bmatrix} x_t \\ e_t \end{bmatrix}, \quad \tilde{u}_t \triangleq \begin{bmatrix} u_t \\ 0 \end{bmatrix}, \quad (62)$$

则其最不利模型为

$$\begin{cases} \xi_{t+1} = \tilde{A}_t \xi_t + \tilde{U} \tilde{u}_t + \tilde{B}_t \epsilon_t, \\ y_t = \tilde{C}_t \xi_t + \tilde{D}_t \epsilon_t, \end{cases} \quad (63)$$

其中:

$$\begin{aligned} \tilde{A}_t &:= \begin{bmatrix} A & BF_t \\ 0 & A - G_t C + (B - G_t D) F_t \end{bmatrix}, \\ \tilde{U} &:= \begin{bmatrix} U \\ 0 \end{bmatrix}, \quad \tilde{B}_t := \begin{bmatrix} B \\ B - G_t D \end{bmatrix} L_t, \\ \tilde{C}_t &:= [C \quad DF_t], \quad \tilde{D}_t := DL_t. \end{aligned} \quad (64)$$

值得注意的是, 最不利的模型是由向后递归实现产生的, 如下所示:

$$\begin{cases} W_{t+1}^+ = \Omega_{t+1}^+ + \theta_t H^T H, \\ K_{v_t} = (I - (B - G_t D)^T W_{t+1}^+ (B - G_t D))^{-1}, \\ F_t = K_{v_t} (B - G_t D)^T W_{t+1}^+ (A - G_t C), \\ \Omega_t^+ = (A - G_t C)^T W_{t+1}^+ (A - G_t C) + F_t^T K_{v_t}^{-1} F_t. \end{cases} \quad (65)$$

最后, 本文基于最不利模型, 可求得其扰动的Riccati方程 \tilde{P}_t . 假设初值为 $0.01 \times I$, 且仅提取出其中 $\tilde{P}_t(3, 3)$ 出的值, 即为所估计出的加速度漂移误差的协方差矩阵 \tilde{P}_t^d . 如图2所示, 基于容许误差为 $c = 10^{-2}$ 的最不利模型, 红色实线代表由标准卡尔曼滤波器(Kalman filter, CF)所产生的漂移的估计误差协方差, 蓝色实线代表由本文所提的鲁棒贝叶斯滤波器

(low-rank robust bayesian estimator, LRB)所产生的漂移的估计误差协方差, 通过拐点位置和误差大小不难看出, 无论是收敛速度还是估计精度, 本文所提的低秩鲁棒贝叶斯估计算法的都明显优于标准卡尔曼滤波算法. 接下来, 本文采取基于容许误差为 $c = 2 \times 10^{-3}$ 和 $c = 8 \times 10^{-3}$ 的最不利模型, 如图3-4可以看出, 随着 c 的减少, 模型的扰动减弱, 且在漂移的误差协方差减少. 同时, LRB的性能和KF之间的性能差距也会随着 c 的减少而减弱. 特别地, 若 $c = 0$, 即为标称模型时, 则两者完全一致, 如图5所示.

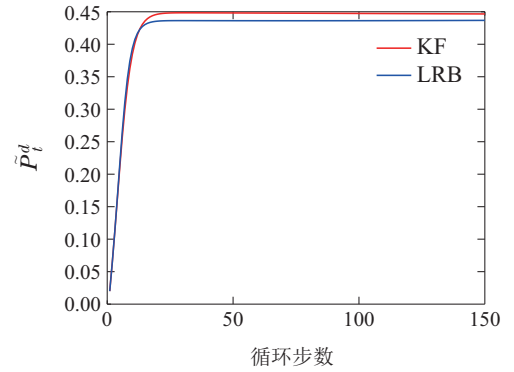


图2 在 $c = 10^{-2}$ 的最不利模型下, KF和LRB的加速度漂移误差协方差

Fig. 2 Variance of the acceleration drift error for KF and LRB under the least favorable model with $c = 10^{-2}$

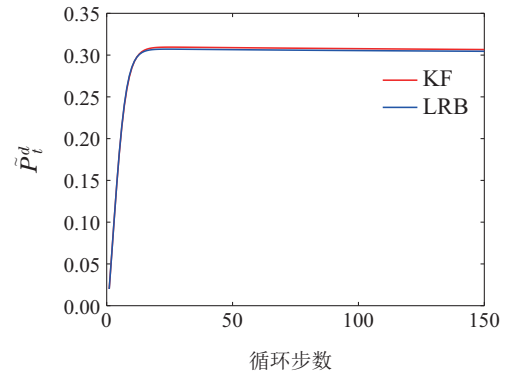


图3 在 $c = 2 \times 10^{-3}$ 的扰动模型下, KF和LRB的加速度漂移误差协方差

Fig. 3 Variance of the acceleration drift error for KF and LRB under the least favorable model with $c = 2 \times 10^{-3}$

5 结论

本文针对不可达系统, 提出了一种在模型扰动下的低秩滤波贝叶斯估计方法. 重要的是, 本文提供了一种全新的证明思路: 首先, 将带KL散度约束的估计问题转换成了一个无约束的拉格朗日函数, 且通过对该函数的二阶变分分别找到了使其在均值和奇异的方差矩阵上为严格凹函数的条件. 其次, 通过求解拉格朗日函数的一阶变分等于0, 得到其均值和方差的最大值. 最后, 证明存在一个唯一的拉格朗日乘子满足其约束条件. 此外, 本文将所提算法应用到MEMS

加速度计漂移校准的应用场景中. 仿真结果证明, 所提算法适用于不可达系统, 且在含有模型扰动情况下, 算法性能高于标准卡尔曼滤波器.

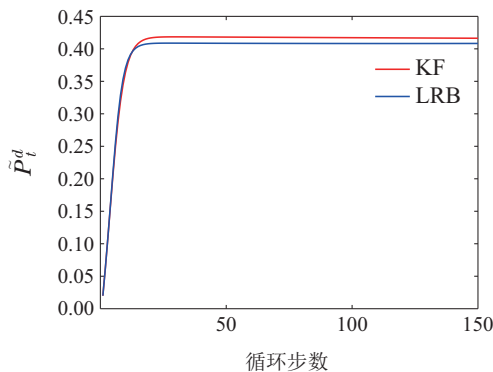


图4 在 $c = 8 \times 10^{-3}$ 的扰动模型下, KF和LRB的加速度漂移误差协方差

Fig. 4 Variance of the acceleration drift error for KF and LRB under the least favorable model with $c = 8 \times 10^{-3}$

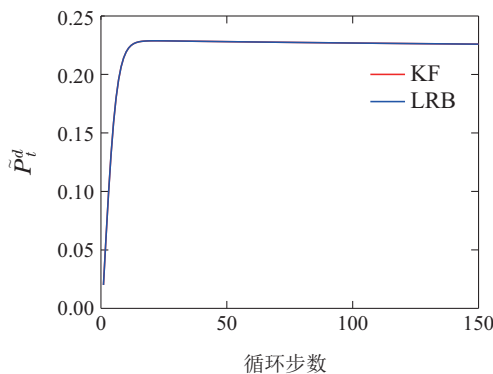


图5 在标称模型下KF和LRB的加速度漂移误差协方差

Fig. 5 Variance of the acceleration drift error for KF and LRB under the nominal model

参考文献:

- [1] ZHU H, ZHANG G, LI Y, et al. A novel robust Kalman filter with unknown non-stationary heavy-tailed noise. *Automatica*, 2021, 127: 109511.
- [2] ALONGE F D, IPPOLITO F, SFERLAZZA A. Sensorless control of induction-motor drive based on robust Kalman filter and adaptive speed estimation. *IEEE Transactions on Industrial Electronics*, 2013, 61(3): 1444 – 1453.
- [3] TALLARINI JR T D. Risk-sensitive real business cycles. *Journal of Monetary Economics*, 2000, 45(3): 507 – 532.
- [4] HANSEN L P, SARGENT T J. *Robustness*. Princeton: Princeton University Press, 2011.
- [5] PETERSEN I R, SAVKIN A V. *Robust Kalman Filtering for Signals and Systems with Large Uncertainties*. Birkhauser, Boston, MA: Springer Science & Business Media, 1999.
- [6] FU M, DE SOUZA C, LUO Z Q. Finite-horizon robust Kalman filter design. *IEEE Transactions on Signal Processing*, 2001, 49(9): 2103 – 2112.
- [7] PENG Meikang, GUO Yunhua, WANG Jingdong, et al. Adaptive target tracking algorithm based on robust cubature Kalman filter. *Control Theory & Applications*, 2020, 37(4): 793 – 800. (彭美康, 郭蕴华, 汪敬东, 等. 基于鲁棒容积卡尔曼滤波的自适应目标跟踪算法. 控制理论与应用, 2020, 37(4): 793 – 800.)
- [8] GUO Yunfei, REN Xin. A Gaussian process based extended target tracking method in polar coordinates. *Control Theory & Applications*, 2020, 37(7): 1501 – 1510. (郭云飞, 任昕. 一种极坐标下基于高斯过程的扩展目标跟踪方法. 控制理论与应用, 2020, 37(7): 1501 – 1510.)
- [9] SPEYER J L, FAN C H, BANAVAR R N. Optimal stochastic estimation with exponential cost criteria. *IEEE Conference on Decision and Control*. Tucson, AZ, USA: IEEE, 1992: 2293 – 2299.
- [10] LEVY B, NIKOUKHAH R. Robust least-squares estimation with a relative entropy constraint. *IEEE Transactions on Information Theory*, 2004, 50(1): 89 – 104.
- [11] LEVY B C, ZORZI M. A contraction analysis of the convergence of risk-sensitive filters. *SIAM Journal on Control and Optimization*, 2016, 54(4): 2154 – 2173.
- [12] LEVY B C, NIKOUKHAH R. Robust state space filtering under incremental model perturbations subject to a relative entropy tolerance. *IEEE Transactions on Automatic Control*, 2012, 58(3): 682 – 695.
- [13] ZORZI M. Robust Kalman filtering under model perturbations. *IEEE Transactions on Automatic Control*, 2016, 62(6): 2902 – 2907.
- [14] ZORZI M. Distributed Kalman filtering under model uncertainty. *IEEE Transactions on Control of Network Systems*, 2019, 7(2): 990 – 1001.
- [15] CHENG M, SHI D, CHEN T. Event-triggered smoothing for hidden Markov models: Risk-sensitive and MMSE results. *Automatica*, 2021, 134: 109933.
- [16] CAI D, LIU W, JI L, et al. Bayesian optimization assisted meal bolus decision based on Gaussian processes learning and risk-sensitive control. *Control Engineering Practice*, 2021, 114: 104881.
- [17] YI S L, ZORZI M. Robust kalman filtering under model uncertainty: The case of degenerate densities. *IEEE Transactions on Automatic Control*, 2022, 67(7): 3458 – 3471.
- [18] YI S L, ZORZI M. Low-rank Kalman filtering under model uncertainty. *The 59th IEEE Conference on Decision and Control (CDC)*. Jeju, South Korea: IEEE, 2020: 2930 – 2935.
- [19] HOLBROOK A. Differentiating the pseudo determinant. *Linear Algebra and its Applications*, 2018, 548: 293 – 304.
- [20] HUANG J, SHI D, CHEN T. Distributed robust state estimation for sensor networks: A risk-sensitive approach. *IEEE Conference on Decision and Control (CDC)*. Miami, FL, USA: IEEE, 2018: 6378 – 6383.
- [21] MEI Junjie, GAO Chengfa, LIU Yongsheng, et al. IMU data quality analysis based on android smartphone. *Proceedings of the 12th China Satellite Navigation Annual Conference*. Nanchang, China: Springer, 2021: 1 – 6. (梅俊杰, 高成发, 刘永胜, 等. 基于Android智能手机的IMU数据质量分析. 第十二届中国卫星导航年会论文集. 中国, 南昌: Springer, 2021: 1 – 6.)
- [22] YI S L, SU T L, JIN X B. Improved smartphone-based indoor localization via drift estimation for accelerometer. *International Conference on Unmanned Systems (ICUS)*. Beijing, China: IEEE, 2017, 379 – 383.

作者简介:

易圣伦 博士, 讲师, 目前研究方向为鲁棒估计、信号处理、非线性系统、信息融合, E-mail: shenglun@dei.unipd.it;

任雪梅 博士, 教授, 博士生导师, 目前研究方向神经网络与自适应控制、系统辨识与非线性控制, 多电机驱动控制, E-mail: xmren@bit.edu.cn.