在不准确方差下带随机系数矩阵的卡尔曼滤波稳定性

周振威 [1] 方海涛 [1]

摘 要 针对离散时间线性随机系统, 研究了卡尔曼滤波的 L_2-稳定性问题. 考虑具有这两个特点的系统: 1) 系数矩阵是随机的; 2) 过程噪声、量测噪声、初始估计误差的方差矩阵不准确. 在系数矩阵有界、条件能观测、初始估计误差有界的假设下, 得到了卡尔曼滤波的 L_2-稳定性. 同时, 建立了卡尔曼滤波和状态空间最小二乘的等价性, 并在该等价性基础上得到状态空间最小二乘的状态估计误差 L_2-稳定性. 最后, 数值仿真说明了卡尔曼滤波的有效性.

关键词 状态估计, 稳定性, 卡尔曼滤波, 状态空间最小二乘

引用格式 周振威, 方海涛. 在不准确方差下带随机系数矩阵的卡尔曼滤波稳定性. 自动化学报, 2013, **39**(1): 43−52

DOI 10.3724/SP.J.1004.2013.00043

L_2-stability of Discrete-time Kalman Filter with Random Coefficients under Incorrect Covariance

ZHOU Zhen-Wei[1] FANG Hai-Tao[1]

Abstract This paper studies the L_2-stability of Kalman filter for discrete-time linear stochastic systems. Two main features, i.e., random coefficient matrices and incorrect covariances of process noise, measurement noise and initial value, are emphasized. Under suitable conditions, including boundedness of coefficient matrices, conditional observability and boundedness of initial error and noises, L_2-stability of Kalman filter is achieved. The equivalence between Kalman filter and state-space least squares algorithm is established. Based on this equivalence, L_2-stability of state estimation error by state-space least squares is also obtained. A numerical example is given to demonstrate the effectiveness of Kalman filtering algorithm.

Key words State estimation, stability, Kalman filter, state-space least squares

Citation Zhen-Wei Zhou, Hai-Tao Fang. L_2-stability of discrete-time Kalman filter with random coefficients under incorrect covariance. *Acta Automatica Sinica*, 2013, **39**(1): 43−52

自从卡尔曼滤波在 20 世纪 60 年代被提出后, 无论在理论方面[1−3] 还是应用方面, 如航天[4−5]、多传感器信息融合[6−9]、传感器网络[10] 和分布式滤波[11], 它都得到了广泛的关注. 其中, 一个基本问题是卡尔曼滤波的稳定性或收敛性分析. 在系数矩阵为确定性矩阵和方差矩阵已知的假设下, 在卡尔曼滤波的稳定性和收敛性方面已有许多重要结果[2−3, 12].

但在工程实际中, 确定性的系数矩阵和方差矩阵已知的假设不一定会得到满足. 例如, 捷联惯性导航系统的初始对准问题[13−14] 或标定问题[15] 可以转化为状态估计问题, 从而利用卡尔曼滤波求解. 对于该问题, 其具有如下两个特点: 1) 系数矩阵含有陀螺和加表的量测信息 (含有量测噪声), 因此系数矩阵是随机的但可以量测; 2) 由于捷联惯性导航系统的噪声项含有慢时变的未知参数, 因此该系统的过程噪声和量测噪声的方差精确值难以获得. 鉴于此, 本文将进一步研究具有如上两个特点的卡尔曼滤波稳定性问题.

当噪声方差矩阵的精确值未知、但系统矩阵为确定性矩阵时, 下述三种方法常被用于系统状态的估计. 1) 当系数矩阵和噪声方差矩阵是常数矩阵时, Mehra[16] 对自适应卡尔曼滤波进行了详细讨论, 将 Bayesian 方法、极大似然方法和相关性方法用于估计未知噪声方差矩阵, 这种方法需要较大运算量; 2) Simon[17] 利用 H_∞ 滤波得到系统状态的稳健估计, 其中这种方法是在诱导 2-范数的意义下, 将有界扰动映射为有界状态估计误差. 但是, 由于 H_∞ 滤波的存在性条件往往难以事先验证, 因此在某种程度上限制了它的应用; 3) 在卡尔曼滤波实施中, 一种实用的方法是直接使用不准确的噪声方差矩阵 (通常为其经验值) 作为算法的参数. 这种做法可能会导致卡尔曼滤波发散. 对于线性定常系统, Sangsuk-Iam 等[18] 在能观测性的假设下证明了估计误差的方差矩阵收敛到某个代数 Riccati 方程的唯一正定解, 对于线性时变系统, Sangsuk-Iam 等[18] 在一致能观测条件下证明了估计误差的方差矩阵保持一致有界.

当系数矩阵是随机矩阵且初始估计误差和噪声方差矩阵的精确值已知时, 许多学者对卡尔曼滤波的稳定性进行了研究. 例如, 相关结论[19−20] 给出

收稿日期 2012-03-09 录用日期 2012-06-19
Manuscript received March 9, 2012; accepted June 19, 2012
国家自然科学基金 (61174143) 资助
Supported by National Natural Science Foundation of China (61174143)
本文责任编委 夏元清
Recommended by Associate Editor XIA Yuan-Qing
1. 中国科学院数学与系统科学研究院系统与控制重点实验室 北京 100190
1. Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190
该文的英文版同时发表在 *Acta Automatica Sinica*, vol. 39, no. 1, pp. 43−52, 2013.

了条件高斯过程的卡尔曼滤波, Bougerol[21] 在平稳遍历的假设下研究了卡尔曼滤波的几乎处处稳定性, Wang 等[22] 在条件能观测的假设下研究了随机 Riccati 方程的 L_p-稳定性. 尽管这些结果在理论上非常深刻, 但在工程中初始估计误差和噪声方差矩阵的精确值往往不易获得 (如捷联惯性导航系统), 因此难以将这些理论直接应用到这些工程问题中. 为此, 有必要为这些工程应用提供一定的理论支持.

本文主要有两点贡献: 1) 对于一般线性随机系统, 给出了卡尔曼滤波的稳定性分析. 与已有的研究工作[18, 22] 相比, 我们考虑的是具有随机系数矩阵和不准确方差这两个特点的系统. 在系数矩阵有界、条件能观测和初值估计误差有界的假设下, 本文得到了卡尔曼滤波的 L_2-稳定性; 2) 建立了状态空间最小二乘和卡尔曼滤波的等价性. 为此, 本文给出了卡尔曼滤波的另一种理解方式. 基于这个等价性, 状态空间最小二乘的估计误差也是 L_2-稳定的.

本文的结构如下: 第 1 节给出卡尔曼滤波稳定性问题的描述, 第 2 节给出卡尔曼滤波的 L_2-稳定性分析, 第 3 节建立状态空间最小二乘的稳定性, 第 4 节和第 5 节分别给出数值仿真和结论.

1 问题描述

本节给出卡尔曼滤波稳定性问题的描述. 为此, 引入 L_4-稳定性、L_2-稳定性和 L_4-指数稳定性的定义. 对于更一般稳定性定义, 如 L_p-稳定性和 L_p-指数稳定性, 可参考文献 [22−24].

定义 1. 对于基本概率空间 $(\Omega, \mathscr{F}, P)$ 的矩阵序列 A_k, 如果 $\sup_k \mathrm{E} \parallel A_k \parallel^4 < \infty$, 则称矩阵序列 A_k 是 L_4-稳定的, 其中范数 $\parallel \cdot \parallel$ 定义为矩阵 X 的最大奇异值, 即 $\parallel X \parallel = \lambda_{\max}^{\frac{1}{2}}(XX^{\mathrm{T}})$.

定义 2. 对于基本概率空间 $(\Omega, \mathscr{F}, P)$ 的向量序列 $\boldsymbol{x}_k$, 如果 $\sup_k \mathrm{E}\|\boldsymbol{x}_k\|^2 < \infty$, 则称矩阵序列 $\boldsymbol{x}_k$ 是 L_2-稳定的, 其中范数 $\parallel \cdot \parallel$ 定义为 $\parallel \boldsymbol{X} \parallel = \sqrt{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}$.

定义 3. 如果存在常数 $M > 0$ 和 $\lambda \in (0, 1)$ 使得:

$$\left(\mathrm{E}\left\|\prod_{k=m}^{n} A_k\right\|^4\right)^{\frac{1}{4}} \leq M\lambda^{n-m+1}, \ \ \forall n \geq m > 0$$

成立, 则称方程 $\boldsymbol{x}_{k+1} = A_k\boldsymbol{x}_k$ 是 L_4-指数稳定的.

考虑状态空间形式的离散线性随机系统:

$$\begin{cases} \boldsymbol{x}_{k+1} = A_k\boldsymbol{x}_k + \boldsymbol{w}_k \\ \boldsymbol{y}_k = C_k\boldsymbol{x}_k + \boldsymbol{v}_k \end{cases} \tag{1}$$

其中, $\boldsymbol{w}_k \in \mathbf{R}^n$ 和 $\boldsymbol{v}_k \in \mathbf{R}^m$ 分别是过程噪声和量测噪声, $\boldsymbol{y}_k \in \mathbf{R}^m$ 和 $\boldsymbol{x}_k \in \mathbf{R}^n$ 分别是系统输出和状态, $A_k \in \mathbf{R}^{n\times n}$ 和 $C_k \in \mathbf{R}^{m\times n}$ 是随机矩阵.

对于系统 (1), 预测形式的卡尔曼滤波为

$$\begin{aligned} \hat{\boldsymbol{x}}_{k+1} &= A_k(\hat{\boldsymbol{x}}_k + K_k(\boldsymbol{y}_k - C_k\hat{\boldsymbol{x}}_k)) \\ K_k &= P_kC_k^{\mathrm{T}}(C_kP_kC_k^{\mathrm{T}} + R_k)^{-1} \\ P_{k+1} &= A_k(P_k - P_kC_k^{\mathrm{T}}(C_kP_kC_k^{\mathrm{T}} + R_k)^{-1}\times \\ &\quad C_kP_k)A_k^{\mathrm{T}} + Q_k \end{aligned} \tag{2}$$

其中, $Q_k \geq Q$ 和 $R_k \geq R$, $Q \in \mathbf{R}^{n\times n}, R \in \mathbf{R}^{m\times m}$ 是任意正定矩阵, $\hat{\boldsymbol{x}}_0$ 和 $P_0 > 0$ 可以任意给定. 即 $Q_k, R_k, P_0, \hat{\boldsymbol{x}}_0$ 不一定是真实的过程噪声方差矩阵、量测噪声方差矩阵、初始估计误差方差矩阵、初始误差均值. 在工程中, $Q_k, R_k, P_0, \hat{\boldsymbol{x}}_0$ 通常是根据经验选取.

根据式 (1) 和式 (2), 卡尔曼滤波的估计误差满足如下方程:

$$\tilde{\boldsymbol{x}}_{k+1} = A_k(I - K_kC_k)\tilde{\boldsymbol{x}}_k + \boldsymbol{w}_k - A_kK_k\boldsymbol{v}_k \tag{3}$$

我们考虑的问题是: 在适当的条件下, 给出卡尔曼滤波误差 $\tilde{\boldsymbol{x}}_k$ 的稳定性分析. 其中, 本文不要求系数矩阵 A_k, C_k 具有平稳性和遍历性. 此外, 噪声项 $\boldsymbol{w}_k, \boldsymbol{v}_k$ 可以是相关的随机噪声或者有界扰动.

注 1. 本文考虑的问题和已有工作存在着密切联系. 例如, 当 A_k, C_{k+1} 是 $\mathscr{H}_k = \sigma(\boldsymbol{y}_0, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_k)$ 可测, $\boldsymbol{w}_k, \boldsymbol{v}_k$ 分别是零均值、方差矩阵分别为Q_k, R_k 的高斯白噪声时, 算法 (2) 是条件正态过程的卡尔曼滤波[19−20], 其中 $\sigma(\boldsymbol{y}_0, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_k)$ 是由 $\boldsymbol{y}_0, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_k$ 所生成的 σ-代数. 又如, 当输出为一维且状态矩阵为单位矩阵时, Guo 等[25−27] 给出了卡尔曼滤波稳定性分析. 当 A_k, C_k 为确定性矩阵时, 本文所讨论的卡尔曼滤波稳定性问题就退化为 Sangsuk-Iam 等[18] 所讨论的情形.

注 2. 根据卡尔曼滤波理论, 如果 $\boldsymbol{w}_k(\boldsymbol{v}_k)$ 是零均值、方差为 $Q_k(R_k)$ 的不相关过程 (量测) 噪声, $\boldsymbol{x}_0$ 与 $\boldsymbol{w}_k, \boldsymbol{v}_k$ 互不相关, 那么 $\hat{\boldsymbol{x}}_k$ 是 $\boldsymbol{x}_k$ 的线性无偏最小方差估计. 此外, 在式 (2) 定义的 P_k 是滤波误差的方差矩阵, 即 $P_k = \mathrm{E}((\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}})$. 对于 $P_k = \mathrm{E}((\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}})$ 的收敛性或有界性问题, 这已有许多经典的结果[1−3]. 但在工程中, $\boldsymbol{w}_k(\boldsymbol{v}_k)$ 的精确方差可能未知, 且 $\boldsymbol{w}_k(\boldsymbol{v}_k)$ 是相关噪声或有界噪声. 对于这些情形, 一种可行的状态估计算法是在算法 (2) 中直接采用正定矩阵 $Q_k > Q > 0, R_k > R > 0$ 作为参数. 因此, $\hat{\boldsymbol{x}}_k$ 不一定是 $\boldsymbol{x}_k$ 的线性无偏最小方差估计, $P_k = \mathrm{E}((\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)^{\mathrm{T}})$ 也不一定成立. 在不精确方差下, 本文将重点讨论一般线性随机系统的卡尔曼滤波 L_2-稳定性, 即 $\mathrm{E} \parallel \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k \parallel^2$ 的有界性.

2 卡尔曼滤波的 L_2-稳定性

本节给出算法 (2) 的稳定性.

2.1 几个假设

为给出主要结果, 需要下述几个假设:

A1 (系数矩阵的有界性). 系数矩阵 A_k, C_k 有

界, 即存在两个正常数 M_1, M_2 使得如下两式成立:

$$\sup_k \| A_k \| \le M_1,\ \ \sup_k \| C_k \| \le M_2$$

A2 (条件能观测性). 对于任意 $\varepsilon > 0$, 存在常数 $\delta > 0$ 满足:

$$P\{\lambda_{\min}(O(k+h,k)) > \delta | \mathscr{G}_{k-1}\} > 1-\varepsilon, \forall k$$

h 是某个正整数, $O(k+h,k) = \sum_{i=k}^{k+h-1} \Phi^{\mathrm{T}}(i,k) \times C_i^{\mathrm{T}} C_i \Phi(i,k)$, $\mathscr{G}_k = \sigma(A_i, C_i, i \le k)$ 是由 $(A_i, C_i, i \le k)$ 生成的 σ-代数.

A3 (噪声和初始误差的有界性). 存在正常数 M_3, M_4, M_5 使得 $\mathrm{E} \| \tilde{\boldsymbol{x}}_0 \|^4 \le M_3, \mathrm{E} \| \boldsymbol{w}_k \|^4 \le M_4, \mathrm{E} \| \boldsymbol{v}_k \|^8 \le M_5$.

注 3. 如果 (A_k, C_k) 满足 A2, 我们称对于 $\mathscr{G}_k(A_k, C_k)$ 是条件能观测的. 这个定义的提出是用于保证随机 Riccati 方程的稳定性, 详细讨论可参考文献 [22]. 为验证该条件, 可检验如下更强的条件:

B2 (一致能观测性). 存在常数 δ 和正整数 h 满足: $\lambda_{\min}(O(k+h,k)(\omega)) \ge \delta > 0, \forall k$, 对于几乎每根轨道$\omega$. 即对几乎每根轨道 ω, (A_k, C_k) 都是一致能观测的.

易见, B2 蕴含了 A2, 反之则不一定成立. 下面给出一个一维的例子以说明这点.

例 1. 假定 $A_k = 1$, $C_k = \gamma_k$ 在 $[0,1]$ 上是服从均匀分布的独立同分布序列. 对于任意 $\varepsilon > 0$, 存在常数 $h_0 = 1, \delta = \frac{\varepsilon}{2}$ 满足: $P\{\gamma_k^2 > \delta\} = P\{\gamma_k^2 > \frac{\varepsilon}{2}\} = 1 - \frac{\varepsilon}{2} > 1 - \varepsilon$. 因此, A2 成立. 但是, 对于任意整数 $h > 0$ 和任意 $\delta > 0$, 有:

$$P\{\gamma_k^2 + \gamma_{k+1}^2 + \cdots + \gamma_{k+h-1}^2 \le \delta\} = \int_0^{\delta} \mathrm{d}x_1 \int_0^{\delta - x_1} \mathrm{d}x_2 \cdots \int_0^{\delta-(x_1+x_2+\cdots+x_{h-1})} \mathrm{d}x_h = \frac{\delta^h}{h!} > 0$$

因此, B2 不成立.

注 4. 如果 $\tilde{\boldsymbol{x}}_0, \boldsymbol{w}_k, \boldsymbol{v}_k$ 有界, 即存在常数 M_6, M_7, M_8 满足 $\|\tilde{\boldsymbol{x}}_0\| \le M_6, \|\boldsymbol{w}_k\| \le M_7, \|\boldsymbol{v}_k\| \le M_8$, 则 A3 也成立. 此外, 如果 $\tilde{\boldsymbol{x}}_0, \boldsymbol{w}_k, \boldsymbol{v}_k$ 具有如下形式:

$$\tilde{x}_0 = \tilde{\boldsymbol{x}}_{0,1} + \tilde{\boldsymbol{x}}_{0,2}, \boldsymbol{w}_k = \boldsymbol{w}_{k,1} + \boldsymbol{w}_{k,2}, \boldsymbol{v}_k = \boldsymbol{v}_{k,1} + \boldsymbol{v}_{k,2}$$

其中, $\tilde{\boldsymbol{x}}_{0,1}, \boldsymbol{w}_{k,1}, \boldsymbol{v}_{k,1}$ 满足 A3, $\tilde{\boldsymbol{x}}_{0,2}, \boldsymbol{w}_{k,2}, \boldsymbol{v}_{k,2}$ 分别以 M_6, M_7, M_8 为上界. 由 C_r-不等式易知, $\tilde{\boldsymbol{x}}_0, \boldsymbol{w}_k, \boldsymbol{v}_k$ 也满足 A3. 例如, $\mathrm{E} \| \tilde{\boldsymbol{x}}_0 \|^4 \le 2^3(\mathrm{E} \| \tilde{\boldsymbol{x}}_{0,1} \|^4 + \mathrm{E} \| \tilde{\boldsymbol{x}}_{0,2} \|^4) = 8(M_3 + M_6^4)$.

此外, 任何条件能观测的 (A_k, C_k) 经过有界反馈后仍保持条件能观测. 为此, 有下述命题:

命题 1. 假设 A1 和 A2 成立, 则对于任意一致有界的增益 K_k ($\mathscr{G}_k$-可测), $(A_k - K_k C_k, C_k)$ 仍保持条件能观测 (相对于 $\mathscr{G}_k$).

证明. 见附录 A. □

2.2 估计误差的 L_2-有界性

我们将分三个步骤来分析估计误差 $\tilde{\boldsymbol{x}}_k$ 的 L_2-稳定性.

首先, L_4-有界性可由下述引理得到:

引理 1[22]**.** 假设 A1 和 A2 成立, 则随机 Riccati 差分方程

$$P_{k+1} = A_k(P_k - P_k C_k^{\mathrm{T}}(C_k P_k C_k^{\mathrm{T}} + R_k)^{-1} C_k P_k) A_k^{\mathrm{T}} + Q_k \tag{4}$$

是 L_4-稳定的, 即 $\sup_k \mathrm{E} \| P_k \|^4 < \infty$. 其中, $Q_k \ge Q > 0$, $R_k \ge R > 0$, $P_0 > 0$ 为任意正定矩阵.

其次, 矩阵序列 P_k 具有 L_4-指数稳定性.

将随机 Riccati 方程 (4) 改写为下述 Lyapunov 方程:

$$P_{k+1} = A_k(I - K_k C_k) P_k (I - K_k C_k)^{\mathrm{T}} A_k^{\mathrm{T}} + \bar{Q}_k \tag{5}$$

其中, $K_k = P_k C_k^{\mathrm{T}}(C_k P_k C_k^{\mathrm{T}} + R_k)^{-1}, \bar{Q}_k = A_k K_k \times R_k^{-1} K_k^{\mathrm{T}} A_k^{\mathrm{T}} + Q_k$.

引理 2[22]**.** 假设 A1 和 A2 成立, 则矩阵序列 $\{A_k(I - K_k C_k)\}$ 是 L_4-指数稳定的, 即存在常数 $M > 0$ 和 $\lambda \in (0,1)$ 使得下式成立:

$$\left(\mathrm{E} \left\| \prod_{k=m}^{n} A_k(I - K_k C_k) \right\|^4 \right)^{\frac{1}{4}} \le M \lambda^{n-m+1}$$

最后, 由下述定理可得状态估计误差的 L_2-有界性.

定理 1. 假设 A1, A2 和 A3 成立, 则状态估计均方误差 $\tilde{\boldsymbol{x}}_k = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k$ 保持有界, 即:

$$\sup_k \mathrm{E} \| \tilde{\boldsymbol{x}}_k \|^2 < \infty$$

证明. 由式 (2) 可知:

$$\tilde{\boldsymbol{x}}_{k+1} = \prod_{j=-1}^{k-1} (A_j(I - K_j C_j)) \tilde{\boldsymbol{x}}_0 + \sum_{i=0}^{k} \prod_{j=i-1}^{k-1} (A_j(I - K_j C_j)) \boldsymbol{\delta}_i \tag{6}$$

其中, $\boldsymbol{\delta}_i = \boldsymbol{w}_i - A_i K_i \boldsymbol{v}_i$ 和 $A_{-1}(I - K_{-1} C_{-1}) = I$.

根据 K_k 的定义和迹的性质, 可得:

$$\begin{aligned} \| K_k \|^2 &= \lambda_{\max}(P_k C_k^{\mathrm{T}}(C_k P_k C_k^{\mathrm{T}} + R_k)^{-2} C_k P_k) \le \\ &\mathrm{tr}(P_k C_k^{\mathrm{T}}(C_k P_k C_k^{\mathrm{T}} + R_k)^{-2} C_k P_k) = \\ &\mathrm{tr}((C_k P_k C_k^{\mathrm{T}} + R_k)^{-1} C_k P_k^2 C_k^{\mathrm{T}}(C_k P_k C_k^{\mathrm{T}} + \\ &R_k)^{-1}) \le \lambda_{\max}(P_k) \mathrm{tr}((C_k P_k C_k^{\mathrm{T}} + R_k)^{-1} \times \\ &C_k P_k C_k^{\mathrm{T}}(C_k P_k C_k^{\mathrm{T}} + R_k)^{-1}) = \\ &\lambda_{\max}(P_k) \mathrm{tr}((C_k P_k C_k^{\mathrm{T}} + R_k)^{-1}) \le \\ &\| P_k \| \mathrm{tr}((C_k P_k C_k^{\mathrm{T}} + R_k)^{-1}) \le \\ &\mathrm{tr}(R^{-1}) \| P_k \| \end{aligned} \tag{7}$$

根据式 (6) 和 Minkovski 不等式, 可得:

$$(\mathrm{E}\|\tilde{\boldsymbol{x}}_{k+1}\|^2)^{\frac{1}{2}} = \left(\mathrm{E}\left\|\prod_{j=-1}^{k-1}(A_j(I-K_jC_j))\tilde{\boldsymbol{x}}_0 + \sum_{i=0}^{k}\prod_{j=i-1}^{k-1}(A_j(I-K_jC_j))\boldsymbol{\delta}_i\right\|^2\right)^{\frac{1}{2}} \leq \left(\mathrm{E}\left\|\prod_{j=-1}^{k-1}(A_j(I-K_jC_j))\tilde{\boldsymbol{x}}_0\right\|^2\right)^{\frac{1}{2}} + \sum_{i=0}^{k}\left(\mathrm{E}\left\|\prod_{j=i-1}^{k-1}(A_j(I-K_jC_j))\boldsymbol{\delta}_i\right\|^2\right)^{\frac{1}{2}} := I_1$$

根据 Hölder 不等式, 可得:

$$I_1 \leq \left(\mathrm{E}\left\|\prod_{j=-1}^{k-1}(A_j(I-K_jC_j))\right\|^4\right)^{\frac{1}{4}} \times (\mathrm{E}\parallel\tilde{\boldsymbol{x}}_0\parallel^4)^{\frac{1}{4}} + \sum_{i=0}^{k}\left(\mathrm{E}\left\|\prod_{j=i-1}^{k-1}(A_j(I-K_jC_j))\right\|^4\right)^{\frac{1}{4}}(\mathrm{E}\parallel\boldsymbol{\delta}_i\parallel^4)^{\frac{1}{4}} := I_2$$

此外, 由引理 2 和 A3 可知:

$$I_2 \leq O(1) + O(1)\sum_{i=0}^{k}\lambda^{k-i+1}(\mathrm{E}\parallel\boldsymbol{\delta}_i\parallel^4)^{\frac{1}{4}}$$

其中, 符号 $O(1)$ 表示某个有界量, 且 $O(1)$ 具有如下性质: $O(1)+O(1)=O(1)$, $C\times O(1)=O(1)$ (C 是某个常数). 这些性质便于我们分析卡尔曼滤波的稳定性.

对于 $\boldsymbol{\delta}_i\,(i=0,1,2,\cdots,k)$, 根据 Minkovski 不等式可知:

$$(\mathrm{E}\|\boldsymbol{\delta}_i\|^4)^{\frac{1}{4}} = (\mathrm{E}\|\boldsymbol{w}_i - A_iK_i\boldsymbol{v}_i\|^4)^{\frac{1}{4}} \leq (\mathrm{E}\|\boldsymbol{w}_i\|^4)^{\frac{1}{4}} + (\mathrm{E}\|A_iK_i\boldsymbol{v}_i\|^4)^{\frac{1}{4}} := I_3$$

由式 (7) 和 A1 可知:

$$I_3 \leq (\mathrm{E}\|\boldsymbol{w}_i\|^4)^{\frac{1}{4}} + (\mathrm{E}(\|A_i\|^4\|K_i\|^4\|\boldsymbol{v}_i\|^4))^{\frac{1}{4}} \leq (\mathrm{E}\|\boldsymbol{w}_i\|^4)^{\frac{1}{4}} + O(1)(\mathrm{E}(\|P_i\|^2\|\boldsymbol{v}_i\|^4))^{\frac{1}{4}} := I_4$$

根据 Hölder 不等式、引理 1 和 A3 可得:

$$I_4 \leq (\mathrm{E}\|\boldsymbol{w}_i\|^4)^{\frac{1}{4}} + O(1)(\mathrm{E}\|P_i\|^4)^{\frac{1}{8}} \times \mathrm{E}(\|\boldsymbol{v}_i\|^8)^{\frac{1}{8}} = O(1)$$

因此, $(\mathrm{E}\parallel\tilde{\boldsymbol{x}}_{k+1}\parallel^2)^{\frac{1}{2}} = O(1)+O(1)\sum_{i=0}^{k}\lambda^i = O(1)$ 成立, 即 $\mathrm{E}\parallel\tilde{\boldsymbol{x}}_{k+1}\parallel^2 = O(1)$. □

3 状态空间最小二乘的 L_2-稳定性

对于卡尔曼滤波, 可从递推状态空间最小二乘的角度去理解它, 如 [28−30]. Haykin 等[29] 指出: 如果系数矩阵为确定性的, 同时 P_0, Q_k, R_k 分别为初始误差、过程噪声以及量测噪声的精确方差矩阵, 则状态空间最小二乘的稳定性分析可以转化为卡尔曼滤波的稳定性分析. 当初始值、过程噪声和量测噪声的统计性质未知, 且系数矩阵为随机矩阵时, 状态空间最小二乘是否保持稳定还不清楚. 对于这种情形, 我们更关注它的稳定性. 相比于极大似然算法和最小均方算法, 最小二乘通常不需要噪声的统计假设, 因而最小二乘具有较广适用范围. 对于线性随机系统 (1), 先利用命题 2 建立状态空间最小二乘和卡尔曼滤波的等价性, 然后通过定理 2 给出状态空间最小二乘的 L_2-稳定性分析.

命题 2. 考虑如下带约束的动态优化问题:

$$\min_{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{k+1},\boldsymbol{w}_0,\boldsymbol{w}_1,\cdots,\boldsymbol{w}_k} J_k \quad \text{s.t. } \boldsymbol{x}_{i+1} = A_i\boldsymbol{x}_i + \boldsymbol{w}_i,\ i=1,2,\cdots,k \tag{8}$$

其中, $J_k = \frac{1}{2}\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\|^2_{P_0^{-1}} + \frac{1}{2}\sum_{i=0}^{k}\|\boldsymbol{y}_i - C_i\boldsymbol{x}_i\|^2_{R_i^{-1}} + \frac{1}{2}\sum_{i=0}^{k}\|\boldsymbol{w}_i\|^2_{Q_i^{-1}}$. 基于数据集合 $\{\boldsymbol{y}_i\}_{i=1}^k$, 该动态优化问题的最优解为 $\hat{\boldsymbol{x}}_{0|k},\hat{\boldsymbol{x}}_{1|k},\cdots,\hat{\boldsymbol{x}}_{k+1|k},\hat{\boldsymbol{w}}_{0|k},\hat{\boldsymbol{w}}_{1|k},\cdots,\hat{\boldsymbol{w}}_{k|k}$, 则 $\hat{\boldsymbol{x}}_{k+1|k}$ 和卡尔曼滤波 (2) 的 $\hat{\boldsymbol{x}}_{k+1}$ 具有相同的递推表达式.

证明. 记

$$D_k = \mathrm{diag}\{0, Q_0^{-1}, Q_1^{-1}, \cdots, Q_k^{-1}\}$$
$$W_k = \mathrm{diag}\{P_0^{-1}, R_0^{-1}, R_1^{-1}, \cdots, R_k^{-1}\}$$
$$\Phi(n,m) = \Phi(n,m+1)A_m, \Phi(m,m) = I$$
$$H_k = \begin{bmatrix}\Phi(k,0) & \Phi(k,1) & \cdots & \Phi(k,k)\end{bmatrix}$$
$$\alpha_k = \begin{bmatrix}C_k\Phi(k,0) & C_k\Phi(k,1) & \cdots & C_k\Phi(k,k)\end{bmatrix}$$
$$U_k = \begin{bmatrix}\boldsymbol{x}_0\\ \boldsymbol{w}_0\\ \boldsymbol{w}_1\\ \vdots\\ \boldsymbol{w}_k\end{bmatrix}, Y_k = \begin{bmatrix}\hat{\boldsymbol{x}}_0\\ \boldsymbol{y}_0\\ \boldsymbol{y}_1\\ \vdots\\ \boldsymbol{y}_k\end{bmatrix}, \begin{bmatrix}\boldsymbol{x}_0\\ C_0\boldsymbol{x}_0\\ C_1\boldsymbol{x}_1\\ \vdots\\ C_k\boldsymbol{x}_k\end{bmatrix} := T_k\begin{bmatrix}\boldsymbol{x}_0\\ \boldsymbol{w}_0\\ \boldsymbol{w}_1\\ \vdots\\ \boldsymbol{w}_k\end{bmatrix}$$

其中, T_k 具有如下表达式:

$$\begin{bmatrix} I & 0 & 0 & \cdots & 0 & 0\\ C_0 & 0 & 0 & \cdots & 0 & 0\\ C_1A_0 & C_1 & 0 & \cdots & 0 & 0\\ C_2A_1A_0 & C_2A_1 & C_2 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ C_k\Phi(k,0) & C_k\Phi(k,1) & C_k\Phi(k,2) & \cdots & C_k & 0 \end{bmatrix}.$$

因为 $\boldsymbol{x}_{i+1} = A_i\boldsymbol{x}_i + \boldsymbol{w}_i\,(i=1,2,\cdots,k)$ 仅由 $\boldsymbol{x}_0,\boldsymbol{w}_0,\boldsymbol{w}_1,\cdots,\boldsymbol{w}_k$ 决定, 所以带约束的动态优化问题 (8) 可以转化为下述无约束的动态优化问题, 即

求 $\hat{\boldsymbol{x}}_{0|k}, \hat{\boldsymbol{w}}_{0|k}, \hat{\boldsymbol{w}}_{1|k}, \cdots, \hat{\boldsymbol{w}}_{k|k}$ 使得下述优化准则达到最小:

$$J_k = \frac{1}{2} \parallel T_k \boldsymbol{U}_k - \boldsymbol{Y}_k \parallel_{W_k}^2 + \frac{1}{2} \parallel \boldsymbol{U}_k \parallel_{D_k}^2$$

通过求导数, 可得上述最小二乘问题的唯一解:

$$\hat{\boldsymbol{U}}_{k|k} = (T_k^{\mathrm{T}} W_k T_k + D_k)^{-1} T_k^{\mathrm{T}} W_k \boldsymbol{Y}_k := S_k T_k^{\mathrm{T}} W_k \boldsymbol{Y}_k$$

其中, $S_k = (T_k^{\mathrm{T}} W_k T_k + D_k)^{-1}$.

因此, 得到:

$$\hat{\boldsymbol{x}}_{k+1|k} = H_{k+1} \hat{\boldsymbol{U}}_{k|k}$$

注意到:

$$\boldsymbol{Y}_{k+1} = \begin{bmatrix} \boldsymbol{Y}_k \\ \boldsymbol{y}_{k+1} \end{bmatrix}, W_{k+1} = \begin{bmatrix} W_k & 0 \\ 0 & R_{k+1}^{-1} \end{bmatrix}$$

$$D_{k+1} = \begin{bmatrix} D_k & 0 \\ 0 & Q_{k+1}^{-1} \end{bmatrix}, T_{k+1} = \begin{bmatrix} T_k & 0 \\ \alpha_{k+1} & 0 \end{bmatrix}$$

以及

$$\begin{aligned} S_{k+1} &= (T_{k+1}^{\mathrm{T}} W_{k+1} T_{k+1} + D_{k+1})^{-1} = \\ &\left(\begin{bmatrix} T_k & 0 \\ \alpha_{k+1} & 0 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} W_k & 0 \\ 0 & R_{k+1}^{-1} \end{bmatrix} \begin{bmatrix} T_k & 0 \\ \alpha_{k+1} & 0 \end{bmatrix} + \right. \\ &\left. \begin{bmatrix} D_k & 0 \\ 0 & Q_{k+1}^{-1} \end{bmatrix} \right)^{-1} = \\ &\begin{bmatrix} (T_k^{\mathrm{T}} W_k T_k + D_k + \alpha_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \alpha_{k+1})^{-1} & 0 \\ 0 & Q_{k+1} \end{bmatrix} = \\ &\begin{bmatrix} (S_k^{-1} + \alpha_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \alpha_{k+1})^{-1} & 0 \\ 0 & Q_{k+1} \end{bmatrix} \end{aligned}$$

于是, 可得:

$$\begin{aligned} \hat{\boldsymbol{U}}_{k+1|k+1} &= S_{k+1} T_{k+1}^{\mathrm{T}} W_{k+1} \boldsymbol{Y}_{k+1} = \\ &\begin{bmatrix} (S_k^{-1} + \alpha_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \alpha_{k+1})^{-1} & 0 \\ 0 & Q_{k+1} \end{bmatrix} \times \\ &\begin{bmatrix} T_k & 0 \\ \alpha_{k+1} & 0 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} W_k & 0 \\ 0 & R_{k+1}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{Y}_k \\ \boldsymbol{y}_{k+1} \end{bmatrix} = \\ &\begin{bmatrix} (S_k^{-1} + \alpha_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \alpha_{k+1})^{-1} & 0 \\ 0 & Q_{k+1} \end{bmatrix} \times \\ &\begin{bmatrix} (T_k^{\mathrm{T}} W_k \boldsymbol{Y}_k + \alpha_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \boldsymbol{y}_{k+1}) \\ 0 \end{bmatrix} \end{aligned}$$

令 $P_{k+1} = H_{k+1} S_k H_{k+1}^{\mathrm{T}}$, 由矩阵求逆公式可知:

$$\begin{aligned} &\hat{\boldsymbol{x}}_{k+2|k+1} = H_{k+2} \hat{\boldsymbol{U}}_{k+1|k+1} = \begin{bmatrix} A_{k+1} H_{k+1} & I \end{bmatrix} \times \\ &\begin{bmatrix} (S_k^{-1} + \alpha_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \alpha_{k+1})^{-1} & 0 \\ 0 & Q_{k+1} \end{bmatrix} \times \\ &\begin{bmatrix} (T_k^{\mathrm{T}} W_k \boldsymbol{Y}_k + \alpha_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \boldsymbol{y}_{k+1}) \\ 0 \end{bmatrix} = A_{k+1} H_{k+1} (S_k - \\ &S_k \alpha_{k+1}^{\mathrm{T}} (\alpha_{k+1} S_k \alpha_{k+1}^{\mathrm{T}} + R_{k+1})^{-1} \alpha_{k+1} S_k)(T_k^{\mathrm{T}} W_k \boldsymbol{Y}_k + \\ &\alpha_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \boldsymbol{y}_{k+1}) = A_{k+1} (H_{k+1} S_k T_k^{\mathrm{T}} W_k \boldsymbol{Y}_k + H_{k+1} \times \\ &S_k \alpha_{k+1}^{\mathrm{T}} (\alpha_{k+1} S_k \alpha_{k+1}^{\mathrm{T}} + R_{k+1})^{-1} (\boldsymbol{y}_{k+1} - \alpha_{k+1} S_k T_k^{\mathrm{T}} \times \\ &W_k \boldsymbol{Y}_k)) = A_{k+1} (\hat{\boldsymbol{x}}_{k+1|k} + P_{k+1} C_{k+1}^{\mathrm{T}} (C_{k+1} P_{k+1} \times \\ &C_{k+1}^{\mathrm{T}} + R_{k+1})^{-1} (\boldsymbol{y}_{k+1} - C_{k+1} \hat{\boldsymbol{x}}_{k+1|k})) = A_{k+1} \times \\ &(\hat{\boldsymbol{x}}_{k+1|k} + K_{k+1} (\boldsymbol{y}_{k+1} - C_{k+1} \hat{\boldsymbol{x}}_{k+1|k})) \end{aligned}$$

其中, $K_{k+1} = P_{k+1} C_{k+1}^{\mathrm{T}} (R_{k+1} + C_{k+1} P_{k+1} C_{k+1}^{\mathrm{T}})^{-1}$.

由矩阵求逆公式可得:

$$\begin{aligned} &P_{k+2} = H_{k+2} S_{k+1} H_{k+2}^{\mathrm{T}} = \\ &\begin{bmatrix} A_{k+1} H_{k+1} & I \end{bmatrix} \begin{bmatrix} * & 0 \\ 0 & Q_{k+1} \end{bmatrix} \begin{bmatrix} H_{k+1}^{\mathrm{T}} A_{k+1}^{\mathrm{T}} \\ I \end{bmatrix} = \\ &A_{k+1} (P_{k+1} - P_{k+1} C_{k+1}^{\mathrm{T}} (C_{k+1} P_{k+1} C_{k+1}^{\mathrm{T}} + \\ &R_{k+1})^{-1} C_{k+1} P_{k+1}) A_{k+1}^{\mathrm{T}} + Q_{k+1} \end{aligned}$$

其中, $*$ 表示 $(S_k^{-1} + \alpha_{k+1}^{\mathrm{T}} R_{k+1}^{-1} \alpha_{k+1})^{-1}$.

因此, 带约束的动态优化问题 (8) 的解 $\hat{\boldsymbol{x}}_{k+1|k}$ 和卡尔曼滤波 (2) 具有相同的递推表达式. □

注 5. 需要指出的是, 虽然卡尔曼滤波和状态空间最小二乘的递推形式相同, 但是它们所优化的准则是完全不同的. 其中, 卡尔曼滤波是通过优化均方误差得到的, 通常要求噪声的统计性质 (如均值和方差) 已知, 而状态空间最小二乘 (8) 则是通过最小化加权平方和得到的, 它并不需要噪声的统计性质.

由命题 2 揭示的等价性可知, 只需分析其中一种算法的稳定性. 当系统 (1) 的系数矩阵退化为确定性系数矩阵, 则状态空间最小二乘的稳定性可利用文献 [18] 的相关结果得到. 本文考虑的是带随机系数矩阵的状态空间最小二乘稳定性问题. 此外, 本文也不需要其他条件 (如 A_k 的可逆性, 噪声 $\boldsymbol{w}_k, \boldsymbol{v}_k$ 和初值 $\boldsymbol{x}_0$ 的高斯性) 来保证动态优化问题 (8) 的解和卡尔曼滤波 (2) 具有相同表达式.

根据定理 1 和命题 2, 得到状态空间最小二乘的稳定性:

定理 2. 假设 A1 ~ A3 成立, 则其估计误差 $\bar{\boldsymbol{x}}_{k+1} = \boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1|k}$ 满足:

$$\sup_k \mathrm{E} \parallel \bar{\boldsymbol{x}}_{k+1} \parallel^2 < \infty$$

其中, $\hat{\boldsymbol{x}}_{k+1|k}$ 由状态空间最小二乘式 (8) 给出.

4 数值仿真

在这部分, 我们以一个仿真例子说明在不精确

方差下卡尔曼滤波 (2) 的有效性. 具体而言, 考虑线性随机系统:

$$\begin{cases}\begin{bmatrix}x_{k+1,1}\\x_{k+1,2}\end{bmatrix}=\begin{bmatrix}a_{11}(k) & 1\\0 & a_{22}(k)\end{bmatrix}\begin{bmatrix}x_{k,1}\\x_{k,2}\end{bmatrix}+\boldsymbol{w}_k\\ y_k=\begin{bmatrix}c_1(k) & 0\end{bmatrix}\begin{bmatrix}x_{k,1}\\x_{k,2}\end{bmatrix}+v_k\end{cases}$$

其中, $a_{11}(k), a_{22}(k), c_1(k)$ 是相互独立的随机过程, 且 $a_{11}(k), a_{22}(k), c_1(k)$ 服从均匀分布:

$$\begin{aligned}a_{11}(k)&\sim \mathrm{U}[1.0, 1.1], \text{i.i.d.}\\ a_{22}(k)&\sim \mathrm{U}[1.0, 1.1], \text{i.i.d.}\\ c_1(k)&\sim \mathrm{U}[-1.0, 1.0], \text{i.i.d.}\end{aligned}$$

关于初始值 $\boldsymbol{x}_0$ 和噪声项 $\boldsymbol{w}_k, v_k$, 对下述两种情形进行仿真.

情形 1. 假定 $\boldsymbol{x}_0$ 服从正态分布 $\mathrm{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right)$, $\boldsymbol{w}_k, v_k$ 是具有下述统计性质的高斯白噪声:

$$\mathrm{E}\boldsymbol{w}_k=\begin{bmatrix}0\\0\end{bmatrix}, Q^0=\mathrm{E}(\boldsymbol{w}_k\boldsymbol{w}_k^{\mathrm{T}})=\begin{bmatrix}0.8 & 0\\0 & 1.2\end{bmatrix}$$

$$\mathrm{E}v_k=0, R^0=\mathrm{E}(v_k^2)=0.9, \mathrm{E}(\boldsymbol{w}_k v_k)=\begin{bmatrix}0\\0\end{bmatrix}$$

显然, A1 和 A3 成立. 同时, A2 成立, 但 B2 不成立, 验证过程可参见附录 B.

取仿真时间为 100 秒, 得到状态的三根轨道, 如图 1 所示.

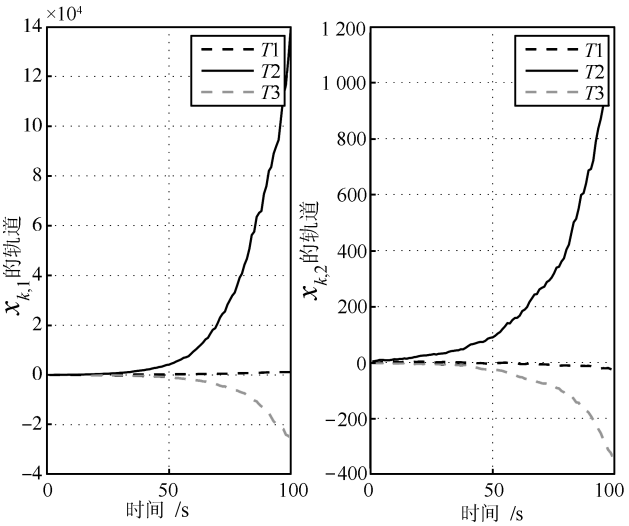

图 1 情形 1 的发散趋势

Fig. 1 Divergence tendencies in Case 1

其中, 在左子图中 $T1$, $T2$, $T3$ 分别表示状态分量 $x_{k,1}$ 的三根不同轨道. 相应地, 右子图的图线也具有类似含义. 图 1 说明 $\boldsymbol{x}_k=\begin{bmatrix}x_{k,1} & x_{k,2}\end{bmatrix}^{\mathrm{T}}$ 随着时间增加趋于无穷, 即该系统是不稳定的.

对于卡尔曼滤波 (2), 分别考虑下述四组参数:

$$1)\ Q_k=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, R_k=1, P_0=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \hat{\boldsymbol{x}}_0=\begin{bmatrix}0\\0\end{bmatrix};$$

$$2)\ Q_k=\begin{bmatrix}8 & 0\\0 & 12\end{bmatrix}, R_k=0.9, P_0=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \hat{\boldsymbol{x}}_0=\begin{bmatrix}0\\0\end{bmatrix};$$

$$3)\ Q_k=\begin{bmatrix}0.8 & 0\\0 & 1.2\end{bmatrix}, R_k=9, P_0=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \hat{\boldsymbol{x}}_0=\begin{bmatrix}0\\0\end{bmatrix};$$

$$4)\ Q_k=\begin{bmatrix}8 & 0\\0 & 12\end{bmatrix}, R_k=9, P_0=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \hat{\boldsymbol{x}}_0=\begin{bmatrix}0\\0\end{bmatrix}.$$

易见, 这四组参数具有相同的 $\hat{\boldsymbol{x}}_0$ 和 P_0, 但它们的噪声方差矩阵 Q_k, R_k 彼此不同. 在第 1 组中, Q_k 与 Q^0, R_k 与 R^0 之差很小. 在第 2 组、第 3 组和第 4 组中, 分别有 $Q_k=10Q^0, R_k=R^0$、$Q_k=Q^0, R_k=10R^0$ 和 $Q_k=10Q^0, R_k=10R^0$. 仿真在 300 根不同轨道上进行, 同时每根轨道的仿真时间均为 100 秒. 对于第一维和第二维状态, 均方估计误差分别以下式进行近似:

$$\begin{aligned}e_{k,1}^{(i)}&=\frac{1}{300}\sum_{j=1}^{300}(x_{k,1}^{(j)}-\hat{x}_{k,1}^{(i,j)})^2\\ e_{k,2}^{(i)}&=\frac{1}{300}\sum_{j=1}^{300}(x_{k,2}^{(j)}-\hat{x}_{k,2}^{(i,j)})^2\end{aligned}\tag{9}$$

其中, $i\,(i=1,2,3,4)$ 分别表示四组不同的参数, $j\,(j=1,2,\cdots,300)$ 分别表示 300 根不同轨道, $k\,(k=1,2,\cdots,100)$ 表示每根轨道的仿真时间, $\boldsymbol{x}_k^{(i,j)}=[x_{k,1}^{(i,j)}x_{k,2}^{(i,j)}]^{\mathrm{T}}$ 是状态的真实值, $\hat{\boldsymbol{x}}_k^{(i,j)}=[\hat{x}_{k,1}^{(i,j)}\hat{x}_{k,2}^{(i,j)}]^{\mathrm{T}}$ 是采用卡尔曼滤波 (2) 得到的估计值. 此外, 记 $e_{k,1}^{(0)}, e_{k,2}^{(0)}$ 为在精确方差矩阵下由卡尔曼滤波得到均方估计误差. 于是, 得到仿真结果如图 2 和图 3 所示. 其中, 图 2 和图 3 的 $e0, e1, e2, e3, e4$ 分别表示 $e_{k,1}^{(0)}, e_{k,1}^{(1)}, e_{k,1}^{(2)}, e_{k,1}^{(3)}, e_{k,1}^{(4)}$ $(e_{k,2}^{(0)}, e_{k,2}^{(1)}, e_{k,2}^{(2)}, e_{k,2}^{(3)}, e_{k,2}^{(4)})$ 的近似值. 图 2 表明在 20 秒后 $e_{k,1}^{(0)}, e_{k,1}^{(1)}, e_{k,1}^{(2)}, e_{k,1}^{(3)}, e_{k,1}^{(4)}$ 分别在 4.0×10^{-2}, 4.0×10^{-2}, 6.0×10^{-2}, $8.0\times10^{-2}, 4.0\times10^{-2}$ 附近变动, $e_{k,1}^{(1)}$、$e_{k,1}^{(4)}$ 和 $e_{k,1}^{(0)}$ 具有相近的性能, $e_{k,1}^{(0)}$ 比 $e_{k,1}^{(2)}$, $e_{k,1}^{(3)}$ 要小. 图 3 表明在 15 秒后 $e_{k,2}^{(0)}, e_{k,2}^{(1)}, e_{k,2}^{(2)}, e_{k,2}^{(3)}, e_{k,2}^{(4)}$ 在 1.5×10^{-2}, 1.5×10^{-2}, 1.8×10^{-2}, 2.0×10^{-2}, 1.5×10^{-2} 附近波动, $e_{k,2}^{(1)}$、$e_{k,2}^{(4)}$ 和 $e_{k,2}^{(0)}$ 的性能相近, $e_{k,2}^{(0)}$ 比 $e_{k,2}^{(2)}, e_{k,2}^{(3)}$ 要小. 因此, 图 2 和图 3 均说明了由卡尔曼滤波 (2) 得到的均方估计误差能保持有界.

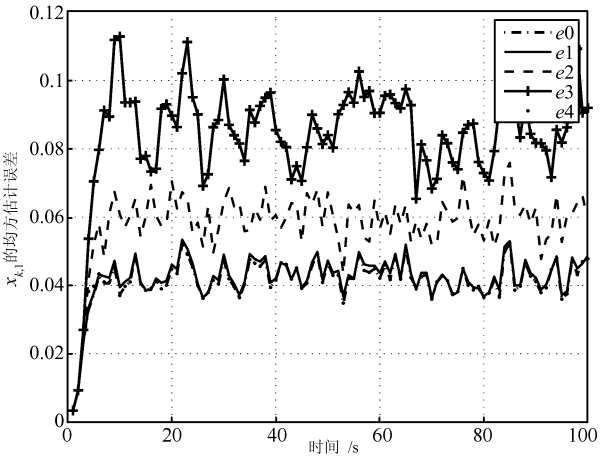

图 2　$x_{k,1}$ 的均方估计误差

Fig. 2　Mean square estimation errors for $x_{k,1}$

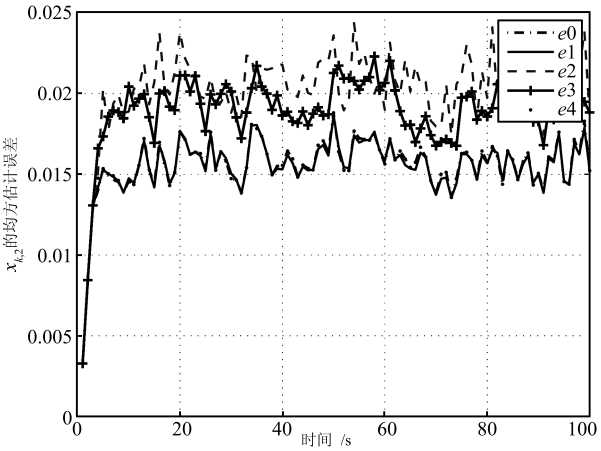

图 3　$x_{k,2}$ 的均方估计误差

Fig. 3　Mean square estimation errors for $x_{k,2}$

情形 2.　假设 $\boldsymbol{x}_0$ 服从正态分布 $\mathrm{N}\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1&0\\0&1\end{bmatrix}\right)$, $\boldsymbol{w}_k, v_k$ 具有如下形式:

$$\boldsymbol{w}_k=\boldsymbol{w}_{k,1}+\boldsymbol{w}_{k,2}, v_k=v_{k,1}+v_{k,2}$$

其中, $\boldsymbol{w}_{k,1}, v_{k,1}$ 为下述高斯白噪声:

$$\mathrm{E}\boldsymbol{w}_{k,1}=\begin{bmatrix}0\\0\end{bmatrix}, Q^0=\mathrm{E}(\boldsymbol{w}_{k,1}\boldsymbol{w}_{k,1}^{\mathrm{T}})=\begin{bmatrix}0.8&0\\0&1.2\end{bmatrix},$$

$$\mathrm{E}v_{k,1}=0, R^0=\mathrm{E}v_{k,1}^2=0.9, \mathrm{E}(\boldsymbol{w}_{k,1}v_{k,1})=\begin{bmatrix}0\\0\end{bmatrix}.$$

此外, $\boldsymbol{w}_{k,2}, v_{k,2}$ 为下述有界扰动:

$$\boldsymbol{w}_{k,2}=\begin{bmatrix}\sin(0.1k)\\\cos(0.1k)\end{bmatrix}, v_{k,2}=\sin(0.1k)$$

与情形 1 类似, 状态分量 $x_{k,1}$ 和 $x_{k,2}$ 的三根状态轨道如图 4 所示. 其中, $T1, T2, T3$ 含义和图 1 的 $T1, T2, T3$ 含义相同. 易见, $\boldsymbol{x}_k=\begin{bmatrix}x_{k,1} & x_{k,2}\end{bmatrix}^{\mathrm{T}}$ 随着时间增加趋于无穷.

此处, 我们在仿真时选取的参数 $\hat{\boldsymbol{x}}_0, P_0, Q_k, R_k$ 与情形 1 相同. 仿真分别在 300 根轨道上进行, 每根轨道的仿真时间为 100 秒, 且均方估计误差由式 (9) 计算. 仿真结果如图 5 和图 6 所示. 其中, $e0, e1, e2, e3, e4$ 的含义和情形 1 的 $e0, e1, e2, e3, e4$ 含义相同. 图 5 说明 $e_{k,1}^{(0)}$, $e_{k,1}^{(1)}$, $e_{k,1}^{(2)}$, $e_{k,1}^{(3)}$, $e_{k,1}^{(4)}$ 分别在 1.0×10^{-1}, 1.0×10^{-1}, 1.0×10^{-1}, 2.5×10^{-1}, 1.0×10^{-1} 附近以三角函数形式 (近似) 周期波动. 同时, 图 6 说明 $e_{k,2}^{(0)}$, $e_{k,2}^{(1)}$, $e_{k,2}^{(2)}$, $e_{k,2}^{(3)}$, $e_{k,2}^{(4)}$ 分别在 4.0×10^{-2}, 4.0×10^{-2}, 4.0×10^{-2}, 5.0×10^{-2}, 4.0×10^{-2} 附近以三角函数形式 (近似) 周期波动. 这两幅图均说明了对于这种形式的过程噪声和量测噪声, 由卡尔曼滤波 (2) 得到的均方估计误差亦能保持有界.

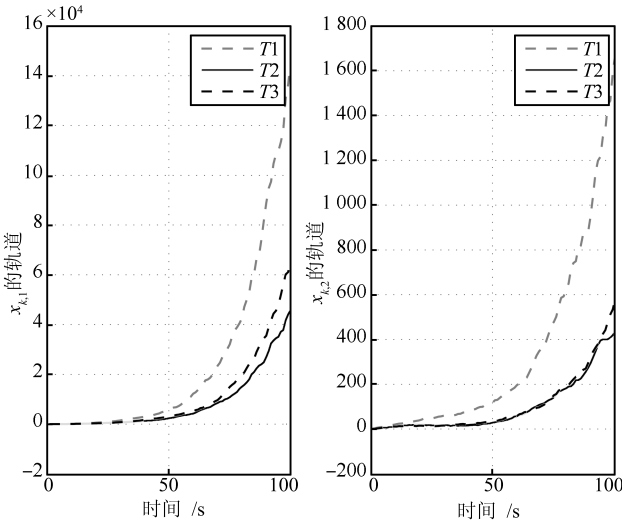

图 4　情形 2 的发散趋势

Fig. 4　Divergence tendency in Case 2

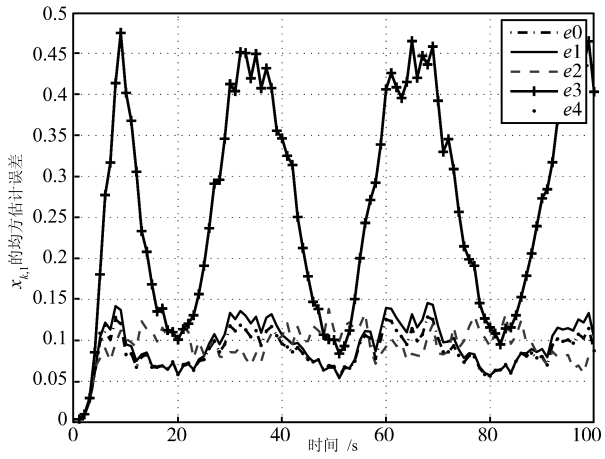

图 5　$x_{k,1}$ 的均方估计误差

Fig. 5　Mean square estimation errors for $x_{k,1}$

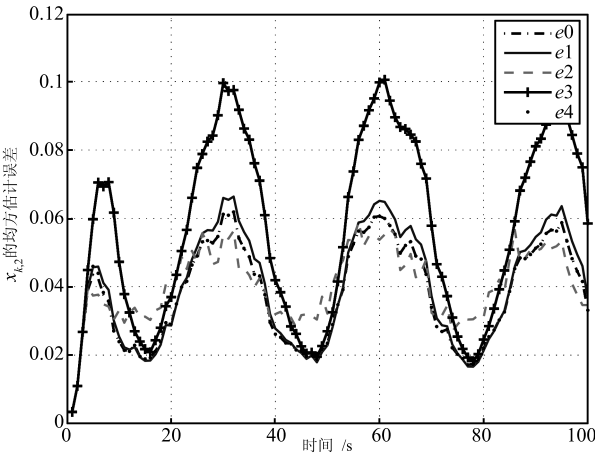

图 6 $x_{k,2}$ 的均方估计误差

Fig. 6 Mean square estimation errors for $x_{k,2}$

5 结论

本文研究了离散时间线性系统的卡尔曼滤波 L_2-稳定性问题. 对于所考虑的系统和算法, 其具有两个特点: 1) 不要求随机系数矩阵满足平稳性和遍历性假设, 这更符合系统控制的要求; 2) 过程噪声和量测噪声可以是非高斯的、相关的或者有界扰动. 因此, 可预期该类系统和算法具有较广的适用范围. 同时, 本文建立了卡尔曼滤波和状态空间最小二乘的等价性, 并在此基础上得到状态空间最小二乘的 L_2-稳定性.

此外, 仍有一些卡尔曼滤波的问题需要解决. 如, 条件能观测 (A2) 是否可放宽就是一个有趣的问题. 从上述仿真可知, 参数设置 (如 Q_k, R_k) 将对卡尔曼滤波的精度产生影响. 因此, 当不能获得噪声 $\boldsymbol{w}_k, \boldsymbol{v}_k$ 的精确统计性质时, 如何选择参数以减小均方估计误差是我们将要研究的重要问题.

附录 A 命题 1 的证明

证 明. 记 $\bar{\Phi}(n,m) = \bar{\Phi}(n,m+1)\bar{A}_m$, 其中, $\bar{A}_k = A_k - K_kC_k$. 于是

$$\begin{aligned}\bar{O}(k+h,k) = &\sum_{i=k}^{k+h-1} \bar{\Phi}^{\mathrm{T}}(i,k)C_i^{\mathrm{T}}C_i\bar{\Phi}(i,k) = \\ &\bar{C}_{k+h,k}\bar{C}_{k+h,k}^{\mathrm{T}} \\ O(k+h,k) = &C_{k+h,k}C_{k+h,k}^{\mathrm{T}}\end{aligned} \tag{A1}$$

其中, $\bar{C}_{k+h,k}, C_{k+h,k}$ 的表达式分别为

$$\begin{aligned}\bar{C}_{k+h,k} = [&C_k^{\mathrm{T}}, \bar{\Phi}^{\mathrm{T}}(k+1,k)C_{k+1}^{\mathrm{T}}, \cdots, \\ &\bar{\Phi}^{\mathrm{T}}(k+h-1,k)C_{k+h-1}^{\mathrm{T}}], \\ C_{k+h,k} = [&C_k^{\mathrm{T}}, \Phi^{\mathrm{T}}(k+1,k)C_{k+1}^{\mathrm{T}}, \cdots, \\ &\Phi^{\mathrm{T}}(k+h-1,k)C_{k+h-1}^{\mathrm{T}}]\end{aligned}$$

另外, $\forall k \le i \le k+h-1$, 有:

$$\begin{aligned}C_i\bar{\Phi}(i,k) = &C_i\bar{\Phi}(i,k+1)(A_k - K_kC_k) = \\ &C_i\bar{\Phi}(i,k+1)A_k - (C_i\bar{\Phi}(i,k+1)K_k)C_k = \\ &C_i\bar{\Phi}(i,k+2)(A_{k+1} - K_{k+1}C_{k+1})A_k - \\ &(C_i\bar{\Phi}(i,k+1)K_k)C_k = C_i\bar{\Phi}(i,k+2)\times \\ &A_{k+1}A_k - (C_i\bar{\Phi}(i,k+2)K_{k+1})C_{k+1}A_k - \\ &(C_i\bar{\Phi}(i,k+1)K_k)C_k\end{aligned}$$

经过迭代计算, 可得:

$$\begin{aligned}C_i\bar{\Phi}(i,k) = &C_i\Phi(i,k) - (C_iK_{i-1})C_{i-1}\times \\ &\Phi(i-1,k) - (C_i\bar{\Phi}(i,i-1)K_{i-2})C_{i-2}\times \\ &\Phi(i-2,k) - \cdots - (C_i\bar{\Phi}(i,k+2)K_{k+1})C_{k+1}\times \\ &\Phi(k+1,k) - (C_i\bar{\Phi}(i,k+1)K_k)C_k := \\ &G_k(i-k+1)C_{k,k+h}^{\mathrm{T}}\end{aligned}$$

其中, $G_k(i-k+1)$ 具有如下形式:

$$\begin{aligned}&[-(C_i\bar{\Phi}(i,k+1)K_k), -(C_i\bar{\Phi}(i,k+2)\times \\ &K_{k+1}), \cdots, -(C_i\bar{\Phi}(i,i-1)K_{i-2}), I, 0, \cdots, 0]\end{aligned}$$

以及 $G_k(i-k+1)$ 的最后 $h-(i-k+1)$ 块是 $m\times m$ 零矩阵.

记 $G_k = [G_k^{\mathrm{T}}(1), G_k^{\mathrm{T}}(2), \cdots, G_k^{\mathrm{T}}(h)]^{\mathrm{T}}$, 于是

$$G_k = \begin{bmatrix} I & 0 & 0 & \cdots & 0 \\ G_k(2,1) & I & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ G_k(h,1) & G_k(h,2) & G_k(h,3) & \cdots & I \end{bmatrix} \tag{A2}$$

其中, $G_k(i,j) = -C_{k+i-1}\bar{\Phi}(k+i-1,k+j)K_{k+j-1}(1 \le j < i \le h)$ 依赖于 $A_i, K_i, C_i, k \le i \le k+h-1$, 且 $G_k(i,j)$ 的对角块是单位矩阵.

因此, 得到:

$$\bar{O}(k+h,k) = C_{k+h,k}G_k^{\mathrm{T}}G_kC_{k+h,k}^{\mathrm{T}} \tag{A3}$$

由式 (A1) 以及 A_k, C_k, K_k 的一致有界性可知: 存在常数 $M_0 > 0$ 使得:

$$\lambda_{\max}(G_k^{\mathrm{T}}G_k) \le M_0$$

记 $G_k^{\mathrm{T}}G_k$ 的特征值分别为 $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_{mh}$, $\lambda_{\max} := \lambda_1$, $\lambda_{\min} := \lambda_{mh}$, 则

$$\prod_{k=1}^{mh}\lambda_k = \det(G_k^{\mathrm{T}}G_k) = \det(G_k^{\mathrm{T}})\det(G_k) = 1 \tag{A4}$$

此外, 有:

$$\prod_{k=1}^{mh}\lambda_k \le \lambda_{\min}(G_k^{\mathrm{T}}G_k)M_0^{mh-1} \tag{A5}$$

由式 (A1), 式 (A3) ~ (A5) 可知:

$$\begin{aligned}\bar{O}(k+h,k) \ge &\lambda_{\min}(G_k^{\mathrm{T}}G_k)C_{k+h,k}C_{k+h,k}^{\mathrm{T}} \ge \\ &\frac{1}{M_0^{mh-1}}O(k+h,k)\end{aligned}$$

因此, $\lambda_{\min}(\bar{O}(k+h,k)) \ge \frac{1}{M_0^{mh-1}}\lambda_{\min}(O(k+h,k))$ 成立. 由 A2 可知, $(A_k - K_kC_k, C_k)$(相对于 σ- 代数 $\mathscr{G}_k$) 也是条件能观测的. □

附录 B 仿真例子的 A2 和 B2 的检验

记 $O(k+2,k)=C_k^{\mathrm{T}}C_k+A_k^{\mathrm{T}}C_{k+1}^{\mathrm{T}}C_{k+1}A_k$, 则

$$\lambda_{\min}(O(k+2,k))=\frac{1}{2}\Big(b_k-\sqrt{b_k^2-4a_k}\Big)=\frac{2a_k}{b_k+\sqrt{b_k^2-4a_k}}\geq\frac{2a_k}{2b_k}=\frac{a_k}{b_k}\geq\frac{a_k}{\bar{b}_k}$$

其中, $b_k=c_1^2(k)+c_1^2(k+1)+a_{11}^2(k)c_1^2(k+1), a_k=c_1^2(k)c_1^2(k+1), a_{11}^2(k)\leq 2, \bar{b}_k=c_1^2(k)+c_1^2(k+1)+2c_1^2(k+1)=c_1^2(k)+3c_1^2(k+1), b_k\leq\bar{b}_k$.

对于任意 $\varepsilon>0$ (不失一般性, 假定 $\varepsilon<0.1$), 根据系数矩阵的独立性可知:

$$\begin{aligned}&P\{\lambda_{\min}(O(k+2,k))\leq\delta\}\leq\\&P\left\{\frac{c_1^2(k)c_1^2(k+1)}{c_1^2(k)+3c_1^2(k+1)}\leq\delta\right\}:=\\&P\left\{\frac{XY}{X+3Y}\leq\delta\right\}=P\{X(Y-\delta)\leq 3\delta Y\}=\\&\int_0^{\delta}\mathrm{d}y\int_0^1\mathrm{d}x+\int_{\delta}^{\frac{\delta}{1-3\delta}}\mathrm{d}y\int_0^1\mathrm{d}x+\\&\int_{\frac{\delta}{1-3\delta}}^1\mathrm{d}y\int_0^{\frac{3\delta y}{y-\delta}}\mathrm{d}x=\delta+\frac{3\delta^2}{1-3\delta}+\\&\left[3\delta\left(1-\frac{\delta}{1-3\delta}\right)+3\delta^2\ln(y-\delta)|_{\frac{\delta}{1-3\delta}}^1\right]\leq\\&\delta+\delta+3\delta+3\delta^2\ln\left(\frac{1-3\delta}{3\delta^2}\right)\leq\delta+\delta+3\delta+\\&3\delta^2\ln\left(\frac{1}{\delta^2}\right)\leq 5\delta+\delta=6\delta\end{aligned}$$

其中, $X=c_1^2(k), Y=c_1^2(k+1)$ 是相互独立的随机变量, 并在 $[0,1]$ 上都服从均匀分布. 因 $\lim_{\delta\to 0^+}\frac{3\delta}{1-3\delta}=0$ 和 $\lim_{\delta\to 0^+}3\delta\ln(\frac{1}{\delta^2})=0$, 故上述第 2 个和第 4 个不等式均成立.

取 $\delta=\frac{1}{6}\varepsilon$, 则 A2 成立. 但是, B2 不再成立. 关于这一点, 从下述不等式得到.

$$\begin{aligned}&\lambda_{\min}(O(k+2,k))=\\&\lambda_{\min}\left(\sum_{i=k}^{k+h-1}\Phi^{\mathrm{T}}(i,k)C_i^{\mathrm{T}}C_i\Phi(i,k)\right)\leq\\&\sum_{i=k}^{k+h-1}\lambda_{\max}(\Phi^{\mathrm{T}}(i,k)C_i^{\mathrm{T}}C_i\Phi(i,k))\leq\\&\sum_{i=k}^{k+h-1}\lambda_{\max}(C_i\Phi(i,k)\Phi^{\mathrm{T}}(i,k)C_i^{\mathrm{T}})\leq\\&\sum_{i=k}^{k+h-1}\lambda_{\max}(M_1^{2(i-k)}C_iC_i^{\mathrm{T}})=\\&\sum_{i=k}^{k+h-1}M_1^{2(i-k)}\lambda_{\max}(C_iC_i^{\mathrm{T}})\end{aligned}$$

其中, 因为 $\|A_k\|=\lambda_{\max}^{\frac{1}{2}}(A_kA_k^T)\leq M_1$, 所以上式最后一个不等式成立.

由于 $c_1^2(i)$ 是相互独立的随机变量, 且在 $[0,1]$ 上都服从均匀分布, 有:

$$\begin{aligned}&P\{\lambda_{\min}(O(k+2,k))\leq\delta\}\geq\\&P\left\{\sum_{i=k}^{k+h-1}M_1^{2(i-k)}\lambda_{\max}(C_iC_i^{\mathrm{T}})\leq\delta\right\}=\\&P\left\{\sum_{i=k}^{k+h-1}M_1^{2(i-k)}c_1^2(i)\leq\delta\right\}=\int_0^{\delta}\mathrm{d}x_1\times\\&\int_0^{\frac{\delta-x_1}{M_1^2}}\mathrm{d}x_2\cdots\int_0^{\frac{\delta-M_1^2x_1-\cdots-M_1^{2(h-2)}x_{h-1}}{M_1^{2(h-1)}}}\mathrm{d}x_h>0\end{aligned}$$

因此, 对于任意 $\delta>0$ 和任意 $h>0$, 都存在一个正的概率集 $\{\lambda_{\min}(O(k+2,k))\leq\delta\}$ 使得 $\lambda_{\min}(O(k+2,k))\leq\delta$, 即 B2 不再成立.

References

1 Jazwinski A H. *Stochastic Processes and Filtering Theory*. New York: Academic Press, 1970. 194−258

2 Anderson B D O, Moore J B. *Optimal Filtering*. New Jersey: Prentice Hall, 1979. 36−128

3 Kailath T, Sayed A H, Hassibi B. *Linear Estimation*. New Jersey: Prentice Hall, 2000. 310−361

4 Grewal M S, Henderson V, Miyasako R. Application of Kalman filtering to the calibration and alignment of inertial navigation systems. *IEEE Transactions on Automatic Control*, 1991, **36**(1): 4−14

5 Grewal M S, Andrews A P. Applications of Kalman filtering in aerospace 1960 to the present−historical perspectives. *IEEE Control Systems Magazine*, 2010, **30**(3): 69−78

6 Luo Y T, Zhu Y M, Luo D D, Zhou J, Song E B, Wang D H. Globally optimal multisensor distributed random parameter matrices Kalman filtering fusion with applications. *Sensors*, 2008, **8**(12): 8086−8103

7 Gu Q T, Fang J. Global optimality for generalized federated filter. *Acta Automatica Sinica*, 2009, **35**(10): 1310−1316

8 Sun S L, Deng Z L. Multi-sensor optimal information fusion Kalman filter. *Automatica*, 2004, **40**(6): 1017−1023

9 Deng Z L, Gao Y, Mao L, Li Y, Hao G. New approach to information fusion steady-state Kalman filtering. *Automatica*, 2005, **41**(10): 1695−1707

10 Sinopoli B, Schenato L, Franceschetti M, Poolla K, Jordan M I, Sastry S S. Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, 2004, **49**(9): 1453−1464

11 Wang S, Yang W, Shi H B. Consensus-based filtering algorithm with packet-dropping. *Acta Automatica Sinica*, 2010, **36**(12): 1689−1696

12 Anderson B D O, Moore J B. Detectability and stabilizability of time-varying discrete-time linear systems. *SIAM Journal on Control and Optimization*, 1981, **19**(1): 20−32

13 Chung D Y, Lee J G, Park C G, Park H W. Strapdown INS error model for multiposition aligment. *IEEE Transactions on Aerospace and Electronic Systems*, 1996, **32**(4): 1362−1366

14 Lee W K, Lee J G, Roh Y K, Park C G. Modeling quaternion errors in SDINS: computer frame approach. *IEEE Transactions on Aerospace and Electronic Systems*, 1998, **34**(1): 289−299

15 Yang X X. Systematic Calibration Method for Laser Gyro Strapdown Inertial Navigation Systems in Outer Field [Ph. D. dissertation]. *University of Chinese Academy of Sciences*, China, 2010

16 Mehra R K. Approaches to adaptive filtering. *IEEE Transactions on Automatic Control*, 1972, **7**(5): 693−698

17 Simon D. *Optimal State Estimation−Kalman, H_∞ and Nonlinear Approaches*. New Jersey: John Wiley and Sons, 2006. 333−369

18 Sangsuk-Iam S, Bullock T E. Analysis of discrete-time Kalman filtering under incorrect noise covariances. *IEEE Transactions on Automatic Control*, 1990, **35**(12): 1034−1038

19 Lipster R S, Shiryayev A N. *Statistics of Random Processes II*. New York: Springer, 1977. 61−69

20 Chen H F, Guo L. *Identification and Stochastic Adaptive Control*. Boston: Birkhauser, 1991. 54−61

21 Bougerol P. Kalman filtering with random coefficients and contractions. *SIAM Journal on Control and Optimization*, 1993, **31**(4): 942−959

22 Wang Y, Guo L. On stability of random Riccati equations. *Science in China (E), Technological Sciences*, 1999, **42**(2): 136−148

23 Guo L, Ljung L. Exponential stability of general tracking algorithms. *IEEE Transactions on Automatic Control*, 1995, **40**(8): 1376−1387

24 Guo L, Ljung L. Performance analysis of general tracking algorithms. *IEEE Transactions on Automatic Control*, 1995, **40**(8): 1388−1402

25 Guo L. Estimation time-varying parameters by the Kalman filter based algorithm: stability and convergence. *IEEE Transactions on Automatic Control*, 1990, **35**(2): 141−147

26 Zhang J F, Guo L, Chen H F. L_p-stability of estimation errors of Kalman filter for tracking time-varying parameters. *International Journal of Adaptive Control and Signal Processing*, 1991, **5**(3): 155−174

27 Guo L. Stability of recursive stochastic tracking algorithms. *SIAM Journal on Control and Optimization*, 1994, **32**(5): 1195−1225

28 Sayed A H, Kailath T. A state-space approach to adaptive RLS filtering. *IEEE Transactions on Signal Processing Magzine*, 1994, **11**(3): 18−60

29 Haykin S, Sayed A H, Zeidler J, Yee P, Wei P. Adaptive tracking of linear time-variant systems by extended RLS algorithms. *IEEE Transactions on Signal Processing*, 1997, **45**(5): 1118−1128

30 Humpherys J, West J. Kalman filtering with Newton′s method. *IEEE Control Systems Magazine*, 2010, **30**(6): 101−106

周振威 2006 年获中国石油大学 (华东) 信息与计算科学专业学士学位. 2012 年获中国科学院大学运筹学与控制论专业博士学位. 主要研究方向为电子产品可靠性, 离散时间随机系统的辨识和滤波. 本文通信作者.
E-mail: zwzhou@amss.ac.cn
(**ZHOU Zhen-Wei** Received his B. S. degree in information and computing sciences from Petroleum University of China in 2006 and his Ph. D. degree in operations research and control theory from University of Chinese Academy of Sciences in 2012. His research interest covers reliability in electronic products, identification and filtering theory in discrete-time stochastic systems. Corresponding author of this paper.)

方海涛 1990 年获北京大学概率统计专业学士学位, 1993 年获清华大学应用数学专业硕士学位, 1996 年获北京大学应用数学专业博士学位. 1996 年至 1998 年, 在中国科学院系统科学研究所从事博士后研究, 并于 1998 年被聘为该所助理研究员. 现为中国科学院数学与系统科学研究院系统控制重点实验室研究员. 主要研究方向为系统辨识, 随机优化和系统控制. E-mail: htfang@iss.ac.cn
(**FANG Hai-Tao** Received his B. S. degree in probability and statistics and M. S. degree in applied mathematics from Peking University and Tsinghua University in 1990 and 1993, respectively. He received his Ph. D. degree in applied mathematics from Peking University in 1996. From 1996 ∼ 1998, he was a postdoctor at the Institute of Systems Science, Chinese Academy of Sciences (CAS), and joined the Institute as an assistant professor in 1998. He now is with the Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, CAS as a full professor. His research interest covers system identification, stochastic optimization, and systems and control.)