

渐进扩展卡尔曼滤波器

张宏欣,周穗华,冯士民

(海军工程大学兵器工程系,湖北武汉 430033)

摘 要: 渐进贝叶斯方法将贝叶斯更新步骤等效为伪时间上的连续演化过程,以实现状态的后验估计. 本文基于渐进贝叶斯框架,导出一种新的高斯型非线性滤波算法. 在线性高斯条件下推导了渐进贝叶斯方法的精确解;证明了对于由线性高斯解确定的动态系统,其均值和协方差矩阵满足的微分方程与常数状态估计的 Kalman-Bucy 滤波器是一致的. 对于非线性系统,利用一阶 Taylor 展开推导了近似解表达式,进而导出渐进扩展卡尔曼滤波器. 仿真算例表明新滤波器性能较扩展卡尔曼滤波器有大幅提高,且避免了窄形似然函数带来的滤波性能恶化问题.

关键词: 非线性滤波; 渐进贝叶斯; Kalman-Bucy 滤波器; 扩展卡尔曼滤波器

中图分类号: TP202 **文献标识码:** A **文章编号:** 0372-2112 (2017)01-0213-07

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2017.01.029

Progressive Extended Kalman Filter

ZHANG Hong-xin, ZHOU Sui-hua, FENG Shi-min

(Department of Weapon Engineering, Naval University of Engineering, Wuhan, Hubei 430033, China)

Abstract: Progressive Bayesian methods formulate the Bayesian update as continuously pseudo-time probability density evolution to perform posterior state estimation. In this paper we derive a novel Gaussian nonlinear filter based on progressive Bayesian framework. A progressive Bayesian solution is firstly derived under linear Gaussian condition. It is proved that the moment evolution of the dynamic system determined by linear Gaussian solution possess the consistency with Kalman-Bucy filter for constant state estimation. For nonlinear system, by using first order Taylor expansion, an approximate solution is derived and the resultant progressive extended Kalman filter is presented. Simulation results demonstrate the superior performance of progressive extended Kalman filter over extended Kalman filter, moreover the performance degrading of nonlinear filtering caused by narrow shape likelihood is avoided.

Key words: nonlinear filtering; progressive Bayesian; Kalman-Bucy filter; extended Kalman filter

1 引言

在一般滤波估值应用中,常需要利用多个传感器的量测值进行状态估计,如果不同(类型)传感器的相对误差是差异显著的^[1],此时似然函数在状态空间中呈窄形分布^[2]. 由于观测模型的非线性影响,估计状态误差在误差较小观测对应的状态空间投影方向上急剧收敛^[3]. 对于典型的高斯假设滤波算法,如扩展卡尔曼滤波器(Extended Kalman Filter, EKF),无迹卡尔曼滤波器^[4](Unscented Kalman Filter, UKF)以及容积卡尔曼滤波器^[5],这将使得状态协方差矩阵是欠估计的(underestimated). 即协方差矩阵与均方误差矩阵之差非正定. 欠估计会造成滤波器精度下降,稳定性变差甚至出现发散现象;而对于粒子型滤波器则由于后验分布与

先验分布重叠较小,使得重采样效率降低而导致粒子贫化问题^[6,7].

渐进贝叶斯(Progressive Bayesian)将贝叶斯法则表达为连续概率演化过程,为解决前述问题提供了理论依据. 通过引入伪时间变量 λ ,根据贝叶斯法则定义同伦(Homotopy)函数

$$q(\mathbf{x}, \lambda) = g(\mathbf{x}) l^\lambda(\mathbf{x}), \lambda \in [0, 1] \quad (1)$$

其中 \mathbf{x} 为系统状态, $g(\mathbf{x})$ 为先验概率密度, $l(\mathbf{x})$ 为似然函数. 同伦函数 $q(\mathbf{x}, \lambda)$ 定义了当 λ 从0到1连续变化时,从先验分布($\lambda=0$)到(未归一化)后验分布($\lambda=1$)的演化过程中的概率分布. 对 λ 进行离散化,式(1)可重写为

$$q(\mathbf{x}, \lambda_i) = q(\mathbf{x}, \lambda_{i-1}) l^{\delta\lambda}(\mathbf{x}), i = 1, 2, \dots, \Lambda \quad (2)$$

其中, $\delta\lambda = 1/\Lambda$. 由式(2)可知,观测信息是随 λ 变化而

逐渐引入的,且 $l^{\delta\lambda}(\mathbf{x})$ 相比似然函数 $l(\mathbf{x})$ 更为平坦,从而在贝叶斯更新过程中避免了窄形似然函数的出现.

将系统状态 \mathbf{x} 从先验分布到后验分布的演化描述为如下的一阶微分方程,

$$\frac{d\mathbf{x}}{d\lambda} = \mathbf{f}(\mathbf{x}, \lambda) \quad (3)$$

若求得(伪)时间导数 $\mathbf{f}(\mathbf{x}, \lambda)$, 则后验状态统计量可通过数值积分方法求解 $\mathbf{x}|_{\lambda=1}$ 得到.

Daum 等首先采用对数形式的归一化同伦函数,根据 Fokker-Planck 方程(FPE)将问题转化为偏微分方程(PDE)求解.由于观测模型的非线性,目前基于该框架算法均采用粒子实现,即所谓的粒子流滤波器[8-13](Particle Flow Filter, PFF).根据笔者掌握的文献来看, PFF 的求解通常比较复杂,方法繁多[11],目前仍未有广泛应用的稳定算法,且对每个粒子均需进行数值积分,计算代价较高.

本文在高斯假设基础上,借助 Daum 框架,导出一种无需粒子实现的渐进贝叶斯滤波算法.在线性高斯条件下推导了 $\mathbf{f}(\mathbf{x}, \lambda)$ 的解析解,证明了式(3)系统均值和协方差矩阵满足的微分方程与 Kalman-Bucy 滤波器具有一致的形式.利用线性系统所得结论,采用一阶 Taylor 展开推导了 $\mathbf{f}(\mathbf{x}, \lambda)$ 的非线性近似解表达式.该算法在形式上类似于扩展卡尔曼滤波器(Extended Kalman Filter, EKF),故称为渐进卡尔曼滤波器(Progressive EKF, PEKF).通过若干仿真算例,将 PEKF 与经典非线性卡尔曼滤波算法的性能进行了比较.结果表明,相比 EKF, PEKF 滤波精度有大幅提高,在窄形似然函数条件下的稳定性优于经典非线性卡尔曼滤波算法.且保留了 EKF 易于编程实现的优点.

2 问题描述

考虑如下随机差分方程表示的非线性系统状态空间模型:

$$\mathbf{x}_k = \mathbf{a}(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} \quad (4)$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \quad (5)$$

其中, $\mathbf{x}_k \in \mathbb{R}^d$ 为 k 时刻的 d 维状态向量, $\mathbf{y}_k \in \mathbb{R}^m$ 为 m 维观测向量;函数 $\mathbf{a}: \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\mathbf{h}: \mathbb{R}^d \rightarrow \mathbb{R}^m$; $\mathbf{w}_{k-1} \in \mathbb{R}^d$ 为 k 时刻状态噪声向量, $\mathbf{v}_k \in \mathbb{R}^m$ 为观测噪声向量, $\mathbf{w}_{k-1} \sim N(\mathbf{0}, \mathbf{Q}_{k-1})$, $\mathbf{v}_k \sim N(\mathbf{0}, \mathbf{R}_k)$, \mathbf{Q}_{k-1} 为过程噪声协方差矩阵, \mathbf{R}_k 为观测噪声协方差矩阵. $\{\mathbf{w}_{k-1}\}$ 和 $\{\mathbf{v}_k\}$ 相互独立.

k 时刻状态后验概率密度函数可表示为:

$$p(\mathbf{x}_k | \mathbf{Y}_k) = \frac{p(\mathbf{y}_k | \mathbf{x}_k)p(\mathbf{x}_k | \mathbf{Y}_{k-1})}{p(\mathbf{y}_k | \mathbf{Y}_{k-1})} \quad (6)$$

其中, $\mathbf{Y}_k := \{\mathbf{y}_{1:k}\}$ 为 k 时刻历史观测值, $p(\mathbf{y}_k | \mathbf{Y}_{k-1}) = \int_{\mathbf{x}_k} p(\mathbf{y}_k | \mathbf{x}_k)p(\mathbf{x}_k | \mathbf{Y}_{k-1})$ 为与状态无关的归一化常数.

记似然函数 $h(\mathbf{x}) = p(\mathbf{y}_k | \mathbf{x}_k)$, 先验概率密度 $g(\mathbf{x}) = p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$, 由式(6)构造如下归一化的同伦函数:

$$q(\mathbf{x}, \lambda) = \frac{g(\mathbf{x})h^\lambda(\mathbf{x})}{K(\lambda)} \quad (7)$$

其中 $K(\lambda) = \int_{\mathbb{R}^d} g(\mathbf{x})h^\lambda(\mathbf{x})d\mathbf{x}$. 由式(3)可知, $q(\mathbf{x}, \lambda)$ 满足零扩散项(diffusion term)的 Fokker-Planck 方程:

$$\frac{\partial q(\mathbf{x}, \lambda)}{\partial \lambda} = -\nabla_{\mathbf{x}} \cdot (q(\mathbf{x}, \lambda)\mathbf{f}(\mathbf{x}, \lambda)) \quad (8)$$

其中 $\nabla_{\mathbf{x}} \cdot (\cdot)$ 为求散度算子. 对式(7)取自然对数得,

$$\log q(\mathbf{x}, \lambda) = \log g(\mathbf{x}) + \lambda \log h(\mathbf{x}) - \log K(\lambda) \quad (9)$$

对式(9)求关于 λ 的偏导数,利用 $\frac{\partial \log K(\lambda)}{\partial \lambda} = E[\log h(\mathbf{x})]$, 并结合式(8)可得:

$$(\log h(\mathbf{x}) - E[\log h(\mathbf{x})])q(\mathbf{x}, \lambda) = -\nabla_{\mathbf{x}} \cdot (q(\mathbf{x}, \lambda)\mathbf{f}(\mathbf{x}, \lambda)) \quad (10)$$

式(10)将渐进贝叶斯求解转换为 PDE 求解问题. 如何根据 PDE(10)导出适于式(4), (5)滤波问题的渐进贝叶斯解,并以较低的计算复杂度实现是下文所要研究的问题.

3 线性高斯条件下的渐进贝叶斯解

3.1 线性高斯解推导

对于线性系统,根据式(4), (5)中的噪声假设可知,先验概率密度 $g(\mathbf{x})$ 满足高斯分布;似然函数 $h(\mathbf{x})$ 为高斯型,即有

$$g(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d |\mathbf{P}_{k|k-1}|}} e^{-\frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}}_{k|k-1})^T \mathbf{P}_{k|k-1}^{-1} (\mathbf{x} - \hat{\mathbf{x}}_{k|k-1})} \quad (11)$$

$$h(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^m |\mathbf{R}_k|}} e^{-\frac{1}{2}(\mathbf{y}_k - \mathbf{H}_k \mathbf{x})^T \mathbf{R}_k^{-1} (\mathbf{y}_k - \mathbf{H}_k \mathbf{x})} \quad (12)$$

其中 $\hat{\mathbf{x}}_{k|k-1}$ 为 k 时刻的先验估计值, $\mathbf{P}_{k|k-1}$ 为先验协方差矩阵, \mathbf{H}_k 为观测矩阵. 将式(11)和(12)代入式(1), 根据文献[14]中 52 页的推导过程可知同伦函数 $q(\mathbf{x}, \lambda)$ 为高斯分布,

$$q(\mathbf{x}, \lambda) = \frac{1}{\sqrt{(2\pi)^d |\mathbf{P}_{k|k}^\lambda|}} e^{-\frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}}_{k|k}^\lambda)^T (\mathbf{P}_{k|k}^\lambda)^{-1} (\mathbf{x} - \hat{\mathbf{x}}_{k|k}^\lambda)} \quad (13)$$

注意到 $q(\mathbf{x}, \lambda)$ 是 λ 条件下的 \mathbf{x} 的后验概率密度, 即

$$q(\mathbf{x}, \lambda) = p(\mathbf{x} | \mathbf{Y}_k, \lambda), \text{ 记 } E[\cdot] := \int_{\mathbb{R}^d} (\cdot) q(\mathbf{x}, \lambda) d\mathbf{x},$$

定义 λ 条件下的后验均值及协方差矩阵

$$\hat{\mathbf{x}}_{k|k}^\lambda = E[\mathbf{x}],$$

$$\mathbf{P}_{k|k}^\lambda = E[(\mathbf{x} - \hat{\mathbf{x}}_{k|k}^\lambda)(\mathbf{x} - \hat{\mathbf{x}}_{k|k}^\lambda)^T]$$

注意到 $\hat{\mathbf{x}}_{k|k}^\lambda$ 和 $\mathbf{P}_{k|k}^\lambda$ 均为离散观测时刻 k 条件下 $[0, 1]$ 区

间上关于 λ 的连续函数,因此略去离散时刻下标,记 $\hat{\mathbf{x}} = \hat{\mathbf{x}}_{k|k}^\lambda, \mathbf{P}_\lambda = \mathbf{P}_{k|k}^\lambda$.

在线性高斯条件下,设渐进贝叶斯解具有如下形式

$$\mathbf{f}(\mathbf{x}, \lambda) = \mathbf{A}(\lambda)\mathbf{x} + \mathbf{b}(\lambda) \quad (14)$$

其中 $\mathbf{A}(\lambda) \in \mathbb{R}^{d \times d}, \mathbf{b}(\lambda) \in \mathbb{R}^d$. 记 $\mathbf{f} = \mathbf{f}(\mathbf{x}, \lambda)$, 注意到式(10)右端的形式,则可将其重写为

$$\log h(\mathbf{x}) - E[\log h(\mathbf{x})] = -\nabla_{\mathbf{x}} \mathbf{f} - (\nabla_{\mathbf{x}} \log q(\mathbf{x}, \lambda))^T \mathbf{f} \quad (15)$$

将式(12)、(13)和(15)代入式(10)可得

$$\begin{aligned} (\mathbf{x} - \hat{\mathbf{x}}_\lambda)^T \mathbf{P}_\lambda^{-1} \mathbf{f} - \text{tr}(\mathbf{A}(\lambda)) &= (\mathbf{x} - \hat{\mathbf{x}}_\lambda)^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{y}_k - \frac{1}{2} \\ \mathbf{x}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{x} + \frac{1}{2} \text{tr}(\mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{P}_\lambda) &+ \frac{1}{2} \hat{\mathbf{x}}_\lambda^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{x}}_\lambda \end{aligned} \quad (16)$$

上式中利用了 $E[\mathbf{s}^T \mathbf{M} \mathbf{s}] = \text{tr}(\mathbf{M} \mathbf{P}_{ss}) + \mathbf{s}^T \mathbf{M} \mathbf{s}$, 其中 \mathbf{s} 为任意实向量, \mathbf{M} 为相应维度的任意实矩阵, \mathbf{P}_{ss} 为 \mathbf{s} 的协方差矩阵. 对式(16)两边乘以 $(\mathbf{x} - \hat{\mathbf{x}}_\lambda)$, 取期望可得

$$\begin{aligned} E[(\mathbf{x} - \hat{\mathbf{x}}_\lambda)(\mathbf{x} - \hat{\mathbf{x}}_\lambda)^T \mathbf{P}_\lambda^{-1} \mathbf{A}(\lambda) \mathbf{x}] &= \mathbf{A}(\lambda) \hat{\mathbf{x}}_\lambda \quad (17) \\ E[(\mathbf{x} - \hat{\mathbf{x}}_\lambda)(\mathbf{x} - \hat{\mathbf{x}}_\lambda)^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{y}_k] \\ - \frac{1}{2} E[(\mathbf{x} - \hat{\mathbf{x}}_\lambda) \mathbf{x}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{x}] \\ = \mathbf{P}_\lambda \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_\lambda) \end{aligned} \quad (18)$$

可得

$$E(\mathbf{f}) = \mathbf{A}(\lambda) \hat{\mathbf{x}}_\lambda + \mathbf{b}(\lambda) = \mathbf{P}_\lambda \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_\lambda) \quad (19)$$

根据式(19)右端的形式,可设

$$\mathbf{A}(\lambda) = -\alpha \mathbf{P}_\lambda \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \quad (20)$$

$$\mathbf{b}(\lambda) = -(1 - \alpha) \mathbf{P}_\lambda \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{x}}_\lambda + \mathbf{P}_\lambda \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{y}_k \quad (21)$$

将式(20)和(21)代入式(16),可得如下关系

$$\begin{aligned} (\mathbf{x} - \hat{\mathbf{x}}_\lambda)^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{y}_k - \alpha \mathbf{x}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{x} + (2\alpha - 1) \\ \mathbf{x}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{x}}_\lambda + (1 - \alpha) \hat{\mathbf{x}}_\lambda^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{x}}_\lambda + \alpha \text{tr} \\ (\mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{P}_\lambda) = (\mathbf{x} - \hat{\mathbf{x}}_\lambda)^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{y}_k - \frac{1}{2} \mathbf{x}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{x} \\ + \frac{1}{2} \text{tr}(\mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{P}_\lambda) + \frac{1}{2} \hat{\mathbf{x}}_\lambda^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{x}}_\lambda \end{aligned} \quad (22)$$

观察等式两边,显然有 $\alpha = \frac{1}{2}$. 再次利用式(20)和(21),得

$$\mathbf{f}(\mathbf{x}, \lambda) = \mathbf{P}_\lambda \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{y}_k - \frac{\mathbf{H}_k \mathbf{x} + \mathbf{H}_k \hat{\mathbf{x}}_\lambda}{2}) \quad (23)$$

上式为线性高斯条件下的渐进贝叶斯解.

3.2 与 Kalman-Bucy 滤波器解的一致性

考虑状态空间模型描述的连续时间常数估计问题

$$\begin{aligned} \frac{d\mathbf{x}_t}{dt} &= 0 \quad (24) \\ \mathbf{y}_t &= \mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t \end{aligned}$$

其中 \mathbf{x}_t 为连续时间状态向量, \mathbf{H}_t 为观测矩阵, \mathbf{y}_t 为连续

时间观测值. 观测噪声 \mathbf{v}_t 为零均值的高斯过程, $E[\mathbf{v}_t \mathbf{v}_t^T] = \mathbf{R}_t$.

对于式(24), Kalman-Bucy 滤波器^[15]给出了最小均方误差估计解 $\hat{\mathbf{x}}_t$ 及其协方差矩阵 \mathbf{P}_t 满足的微分方程,

$$\begin{aligned} \frac{d\hat{\mathbf{x}}_t}{dt} &= \mathbf{K}_t (\mathbf{y}_t - \mathbf{H}_t \hat{\mathbf{x}}_t) \\ \frac{d\mathbf{P}_t}{dt} &= -\mathbf{K}_t \mathbf{R}_t \mathbf{K}_t^T \\ \mathbf{K}_t &= \mathbf{P}_t \mathbf{H}_t \mathbf{R}_t^{-1} \end{aligned} \quad (25)$$

现在证明,对于式(23)确定的一阶线性随机系统

$$\frac{d\mathbf{x}_\lambda}{d\lambda} = \mathbf{P}_\lambda \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{y}_k - \frac{\mathbf{H}_k \mathbf{x}_\lambda + \mathbf{H}_k \hat{\mathbf{x}}_\lambda}{2}) \quad (26)$$

其中 $\mathbf{x}_\lambda \sim N(\hat{\mathbf{x}}_\lambda, \mathbf{P}_\lambda)$, 状态均值 $\hat{\mathbf{x}}_\lambda$ 和协方差矩阵 \mathbf{P}_λ 满足的微分方程与式(25)具有一致的形式. 由此直接得到式(26)均值满足的微分方程为

$$\frac{d\hat{\mathbf{x}}_\lambda}{d\lambda} = \mathbf{P}_\lambda \mathbf{H}_k^T \mathbf{R}_k^{-1} (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_\lambda) \quad (27)$$

根据协方差矩阵定义,注意到上式中求导与取期望无关,

$$\frac{dE[\mathbf{x}_\lambda \mathbf{x}_\lambda^T]}{d\lambda} = E[\mathbf{f}(\mathbf{x}, \lambda) \mathbf{x}_\lambda^T + \mathbf{x}_\lambda \mathbf{f}(\mathbf{x}, \lambda)^T] \quad (28)$$

$$\frac{d(\hat{\mathbf{x}}_\lambda \hat{\mathbf{x}}_\lambda^T)}{d\lambda} = \frac{d\hat{\mathbf{x}}_\lambda}{d\lambda} \hat{\mathbf{x}}_\lambda^T + \hat{\mathbf{x}}_\lambda \frac{d\hat{\mathbf{x}}_\lambda^T}{d\lambda} \quad (29)$$

将式(23),式(27)分别代入式(28)和式(29),可得

$$\frac{d\mathbf{P}_\lambda}{d\lambda} = -\mathbf{P}_\lambda \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{R}_k \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{P}_\lambda \quad (30)$$

令 $\mathbf{K}_\lambda = \mathbf{P}_\lambda \mathbf{H}_k^T \mathbf{R}_k^{-1}$, 可得到下列表达式:

$$\begin{aligned} \frac{d\hat{\mathbf{x}}_\lambda}{d\lambda} &= \mathbf{K}_\lambda (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_\lambda) \\ \frac{d\mathbf{P}_\lambda}{d\lambda} &= -\mathbf{K}_\lambda \mathbf{R}_k \mathbf{K}_\lambda^T \end{aligned} \quad (31)$$

综上,由式(25)和式(31)可知,在线性高斯条件下,对于式(26)给出的系统,其均值和协方差矩阵满足的微分方程与零过程噪声的常数状态系统的 Kalman-Bucy 滤波器是一致的. 该结论说明,式(23)的一阶线性随机系统给出了线性系统状态空间中满足高斯分布样本从先验分布到后验分布的演化规律,式(31)则给出了相应的均值和协方差矩阵演化规律.

4 渐进扩展卡尔曼滤波器

根据线性高斯条件下解的形式,假设非线性条件下的渐进贝叶斯解具有如下形式:

$$\mathbf{f}(\mathbf{x}, \lambda) = \mathbf{B}(\lambda) \mathbf{h}(\mathbf{x}) + \mathbf{u}(\lambda) \quad (32)$$

其中 $\mathbf{h}(\mathbf{x}) : \mathbb{R}^d \rightarrow \mathbb{R}^m$ 为非线性观测方程. $\mathbf{B}(\lambda) \in \mathbb{R}^{d \times m}, \mathbf{u}(\lambda) \in \mathbb{R}^d$ 为与 \mathbf{x} 无关的项. 将 $\mathbf{h}(\mathbf{x})$ 在 $\hat{\mathbf{x}}_\lambda$ 附近

做一阶泰勒展开,则 $f(\mathbf{x}, \lambda)$ 可近似为

$$f(\mathbf{x}, \lambda) \approx \mathbf{B}(\lambda)(\mathbf{h}(\hat{\mathbf{x}}_\lambda) + \hat{\mathbf{J}}(\hat{\mathbf{x}}_\lambda)\tilde{\mathbf{x}}) + \mathbf{u}(\lambda) \quad (33)$$

其中 $\tilde{\mathbf{x}} = (\mathbf{x} - \hat{\mathbf{x}}) \sim N(0, \mathbf{P}_\lambda)$, $\hat{\mathbf{J}}(\hat{\mathbf{x}}_\lambda) = \left. \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}}$ 为函数 $\mathbf{h}(\mathbf{x})$ 在点 $\hat{\mathbf{x}}$ 处的 Jacobian 矩阵. 将式(32)和式(33)代入式(15),利用第2节中推导过程可得

$$\begin{aligned} \mathbf{B}(\lambda) &= -\frac{1}{2}\mathbf{P}_\lambda \hat{\mathbf{J}}(\hat{\mathbf{x}}_\lambda) \mathbf{R}_k^{-1} \\ \mathbf{u}(\lambda) &= -\frac{1}{2}\mathbf{P}_\lambda \hat{\mathbf{J}}(\hat{\mathbf{x}}_\lambda) \mathbf{R}_k^{-1} \mathbf{h}(\hat{\mathbf{x}}_\lambda) + \mathbf{P}_\lambda \hat{\mathbf{J}}(\hat{\mathbf{x}}_\lambda) \mathbf{R}_k^{-1} \mathbf{y}_k \end{aligned} \quad (34)$$

利用式(33)和(34)即可得到非线性条件下 $f(\mathbf{x}, \lambda)$ 的近似解. 进一步可得非线性条件下, $\hat{\mathbf{x}}_\lambda$ 和 \mathbf{P}_λ 满足

$$\begin{aligned} \frac{d\hat{\mathbf{x}}_\lambda}{d\lambda} &= \mathbf{K}_\lambda (\mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_\lambda)) \\ \frac{d\mathbf{P}_\lambda}{d\lambda} &= -\mathbf{K}_\lambda \mathbf{R}_k \mathbf{K}_\lambda^T \end{aligned} \quad (35)$$

其中 $\mathbf{K}_\lambda = \mathbf{P}_\lambda \hat{\mathbf{J}}(\hat{\mathbf{x}}_\lambda) \mathbf{R}_k^{-1}$, 观察式(35)可知,与2.2小节中证明的结论类似,该式与如下估计问题的扩展 Kalman-Bucy 滤波器^[15]具有一致的形式.

$$\begin{aligned} \frac{d\mathbf{x}_t}{dt} &= 0 \\ \mathbf{y}_t &= \mathbf{h}(\mathbf{x}_t) + \mathbf{v}_t \end{aligned} \quad (36)$$

因此,对于非线性滤波估值问题,可利用式(35)以连续(伪)时间方式进行观测更新. 然而,由于式(35)为连续时间系统,直接利用该式需要很小的数值积分步长,所导致的计算量将难以满足实时要求.

为此,对式(35)进行离散化,假设 $\hat{\mathbf{J}}(\hat{\mathbf{x}}_\lambda) \approx \hat{\mathbf{J}}(\hat{\mathbf{x}}_{\lambda|\lambda-\delta\lambda})$, 根据线性卡尔曼滤波理论^[15],可将 \mathbf{K}_λ 近似为

$$\mathbf{K}_\lambda = \mathbf{P}_{\lambda|\lambda-\delta\lambda} \hat{\mathbf{J}}(\hat{\mathbf{x}}_{\lambda|\lambda-\delta\lambda})^T \mathbf{S}_{\lambda|\lambda-\delta\lambda}^{-1} \quad (37)$$

$$\mathbf{S}_{\lambda|\lambda-\delta\lambda} = \hat{\mathbf{J}}(\hat{\mathbf{x}}_{\lambda|\lambda-\delta\lambda}) \mathbf{P}_{\lambda|\lambda-\delta\lambda} \hat{\mathbf{J}}(\hat{\mathbf{x}}_{\lambda|\lambda-\delta\lambda})^T + \mathbf{R}_k$$

其中 $\delta\lambda$ 为离散化步长. 再根据式(35)推导过程,结合式(36),可得到下列等式

$$\begin{aligned} \hat{\mathbf{x}}_{\lambda|\lambda-\delta\lambda} &= \hat{\mathbf{x}}_{\lambda-\delta\lambda} \\ \mathbf{P}_{\lambda|\lambda-\delta\lambda} &= \mathbf{P}_{\lambda-\delta\lambda} \\ \hat{\mathbf{x}}_\lambda &= \hat{\mathbf{x}}_{\lambda|\lambda-\delta\lambda} + \mathbf{K}_\lambda (\mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_{\lambda|\lambda-\delta\lambda})) \delta\lambda \\ \mathbf{P}_\lambda &= (\mathbf{I} - \mathbf{K}_\lambda \hat{\mathbf{J}}(\hat{\mathbf{x}}_{\lambda|\lambda-\delta\lambda}) \delta\lambda) \mathbf{P}_{\lambda|\lambda-\delta\lambda} \end{aligned} \quad (38)$$

式(37)和(38)即为 PEKF 的观测更新等式. 综上所述,结合 EKF 的时间更新步骤,可得 PEKF 算法步骤如下:

1. $k=0$, 初始化 $\hat{\mathbf{x}}_{0|0}, \mathbf{P}_{0|0}$.
2. $k=1, 2, \dots, T$, T 为观测值序列长度.

时间更新

- (1) 对状态方程进行线性化

$$\mathbf{F}_{k-1} = \left. \frac{\partial \mathbf{a}(\mathbf{x})}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{k|k-1}} \quad (39)$$

- (2) 对状态及协方差矩阵进行时间更新

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1} \mathbf{P}_{k-1|k-1} \mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1} \quad (40)$$

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{a}(\hat{\mathbf{x}}_{k-1|k-1})$$

观测更新

- (1) 初始化:

$$\hat{\mathbf{x}}_\lambda |_{\lambda=0} = \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_\lambda |_{\lambda=0} = \mathbf{P}_{k|k-1}, \Lambda = 1/\delta\lambda$$

- (2) $\lambda = \delta\lambda, 2\delta\lambda, \dots, \Lambda\delta\lambda$

按式(37)和(38)进行渐进更新.

- (3) $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_\lambda |_{\lambda=1}, \mathbf{P}_{k|k} = \mathbf{P}_\lambda |_{\lambda=1}$.

当 $\delta\lambda = 1$ 时, PEKF 算法即退化为 EKF 算法.

5 仿真结果及分析

为验证前述所得结论的合理性以及 PEKF 的有效性,本节首先利用距离-角度观测模型作为演示性算例,说明窄形似然函数对滤波性能的影响. 然后采用单变量非稳定增长模型考察 PEKF 相对于 EKF 的滤波性能提升;最后以高度辅助跟踪模型考察 PEKF 的在不同观测噪声条件下的滤波性能,同时给出 EKF, UKF 和 CKF 的仿真结果作为参考.

除演示算例外,仿真中 PEKF 的渐进更新步长 $\delta\lambda = 0.1$, 均方根误差 (RMSE) 采用如下定义

$$RMSE = \sqrt{\frac{1}{T} \sum_{k=1}^T \text{tr}[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T]} \quad (41)$$

其中 T 为观测序列长度, \mathbf{x}_k 为 k 时刻目标真实状态, $\hat{\mathbf{x}}_{k|k}$ 为目标状态估计值.

5.1 演示算例: 距离-角度观测模型

考虑采用距离-角度观测模型的平面直角坐标系内不动点估计问题^[3],

$$\begin{aligned} \mathbf{x}_k &= \mathbf{x}_{k-1} \\ \mathbf{y}_k &= \begin{pmatrix} \sqrt{\mathbf{x}_{k,1}^2 + \mathbf{x}_{k,2}^2} \\ \text{atan}(\frac{\mathbf{x}_{k,2}}{\mathbf{x}_{k,1}}) \end{pmatrix} + \begin{pmatrix} v_r \\ v_\theta \end{pmatrix} \end{aligned} \quad (42)$$

其中 $\mathbf{x}_k = (\mathbf{x}_{k,1}, \mathbf{x}_{k,2})^T$, 观测噪声 $v_r \sim N(0, \sigma_r^2)$, $v_\theta \sim N(0, \sigma_\theta^2)$, $k=1, 2, \dots, T$.

以上模型状态是恒定的,因此不包含采样时间和过程噪声,从而可以单纯考察观测更新过程对于滤波性能的影响. 采用文献[3]中的参数设置,距离噪声方差 $\sigma_r^2 = 2.5 \times 10^{-5}$, 角度噪声方差 $\sigma_\theta^2 = 6 \times 10^{-3}$. 初始估计值为 $\mathbf{x}_{0|0} = (20, 80)^T$, 状态误差矩阵 $\mathbf{P}_{0|0} = 10^2 \mathbf{I}_{2 \times 2}$, 状态真实位置为 $\mathbf{x}_k = (100, 100)^T$.

图1给出了 EKF, UKF, CKF 以及 PEKF 随滤波迭代次数 k 变化的误差曲线,其中第 k 次迭代误差由下式进行 100 次 MC 计算得到

$$RMSE_k = \sqrt{\frac{1}{100} \sum_{i=1}^{100} \text{tr}[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T]} \quad (43)$$

由图可知,几种算法中 EKF, UKF, CKF 均不同程度

地出现了错误收敛,而本文提出的 PEKF 则收敛到了较小的误差水平,且稳定误差大幅地低于其他算法.文献[3]中已经指出,观测噪声方差的差异使估计状态误差 $P_{k|k}$ 的收敛速度在径向距离分量方向远大于角度分量方向,此时由于非线性近似误差影响,使得卡尔曼增益矩阵不能正确根据角度新息在角度分量方向上对状态进行观测更新,从而导致估计结果只在距离分量上与实际距离误差较小,随着 $P_{k|k}$ 的量值进一步收缩,最终造成错误收敛.而由式(2)可知,渐进更新抑制了观测噪声方差的差异,因而 PEKF 能够避免上述问题.

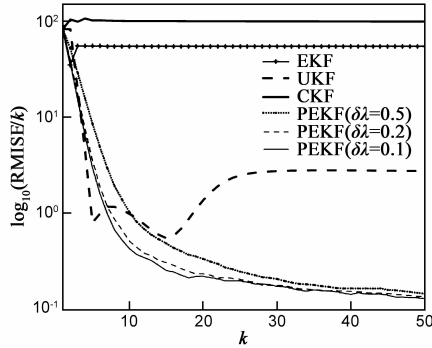


图1 各算法随滤波迭代次数变化的RMSE

图1中同时给出了在不同渐进步长 $\delta\lambda$ 条件下的 PEKF 误差.从图中可知,本例中由于没有引入过程噪声, $\delta\lambda$ 取 0.5 时即可收到较好的效果,此时渐进更新只需要两步迭代.当过程噪声较大时,步长则需进一步减小以提高滤波性能,但过小的步长对于性能提升已十分有限,对于一般应用取 $\delta\lambda = 0.1$ 即可.

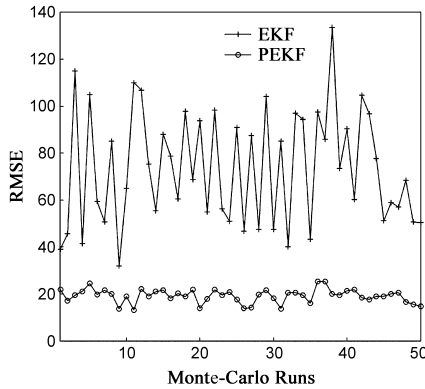


图2 EKF与PEKF经50次MC计算结果($q^2=100$)

5.2 单变量非稳定增长模型

单变量非稳定增长模型是非线性滤波中的一维 Benchmark 问题,其状态空间模型为

$$x_k = \varphi x_{k-1} + \beta \frac{x_{k-1}}{1 + x_{k-1,i}^2} + \gamma \cos(1.2(k-1)) + \omega_{k-1} \quad (44)$$

$$y_k = \frac{x_k^2}{20} + v_k$$

其中模型参数为: $\varphi = 0.5, \beta = 25, \gamma = 8$; 仿真设置为:时刻 $k = 1, 2, \dots, T, T = 150$; 过程噪声 $w_{k-1} \sim N(0, q^2)$, 观测噪声 $v_k \sim N(0, 0.1^2)$. 状态初值 $x_0 = 0.1$. 滤波器初始分布为 $N(x_0, P_{0|0})$, 初始误差矩阵 $P_{0|0} = 0.1$.

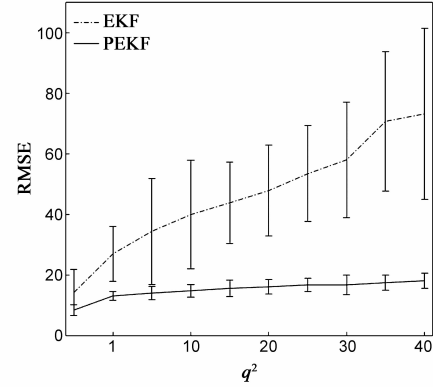


图3 过程噪声增长时RMSE的均值和标准差

图2给出了过程噪声方差 $q^2 = 100$ 时, PEKF 和 EKF 进行 50 次 MC 仿真计算得到的 RMSE, 对比可知 PEKF 在各次仿真的滤波误差均显著低于 EKF; 且误差波动较小, 性能稳定. 图3给出了当过程噪声方差取 $q^2 = 1, 10, 20, \dots, 90$ 时, GPBF 和 PF 算法经 100 次 MC 仿真所得 RMSE 的误差条 (error bar) 曲线, 可知当过程噪声增大时, EKF 的线性化误差急剧增大, 而 PEKF 增长则较为平缓, 且在 100 次 MC 计算后的 RMSE 方差亦小于 EKF. 由于 PEKF 的时间更新步骤与 EKF 完全相同, 因此, 该算例表明 PEKF 的渐进观测更新过程能够减小线性化误差, 相比 EKF 提高了滤波精度和稳定性.

5.3 高度辅助跟踪模型

设目标状态 $x_t = [\rho_t^T, \zeta_t^T]^T$, ρ 为目标的空间直角坐标系位置矢量, ζ 为速度矢量. 假设目标高度 h_k 能被实时测量. 观测矢量为 $y_k = [b_k, r_k, h_k, s_k]^T$, b_k 为观测点 $(0, 0)^T$ 到目标的方位角, r_k 为观测点到目标的距离, s_k 为距离变化率. 相应的观测模型为:

$$b_k = \arctan\left(\frac{\rho_{k,2}}{\rho_{k,1}}\right) + v_{k,1} \quad (45)$$

$$r_k = \sqrt{\rho_{k,1}^2 + \rho_{k,2}^2 + \rho_{k,3}^2} + v_{k,2} \quad (46)$$

$$h_k = \rho_{k,3} + v_{k,3} \quad (47)$$

$$s_k = \frac{\rho_k^T \zeta_k}{r_k} + v_{k,4} \quad (48)$$

其中 v_k 为观测噪声矢量, $v_k \sim N(0, R)$. 目标状态方程为

$$x_k = \Gamma x_{k-1} + w_{k-1}, \Gamma = \begin{pmatrix} I_{3 \times 3} & \Delta t I_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & I_{3 \times 3} \end{pmatrix} \quad (49)$$

过程噪声 $w_k \sim N(\mathbf{0}_{6 \times 1}, Q)$, 观测采样间隔 $\Delta t = 1$, 过程

$$\text{噪声协方差矩阵 } \mathbf{Q} = 5 \cdot \begin{pmatrix} \frac{1}{3} \mathbf{I}_{3 \times 3} & \frac{1}{2} \mathbf{I}_{3 \times 3} \\ \frac{1}{2} \mathbf{I}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{pmatrix}$$

目标初始状态 $\mathbf{x}_0 = [0, 0, 50, 5, 5, 0.1]$, 状态初始分布为 $p_{010}(\mathbf{x}) \sim N(\hat{\mathbf{x}}_0, \mathbf{P}_{010})$, 其中 $\hat{\mathbf{x}}_0 = [2, 2, 50, 6, 6, 0]$, $\mathbf{P}_{010} = \mathbf{I}_{6 \times 6}$. 观测点数 $T = 100$. 为考察滤波器在不同观测误差条件的性能, 设 $\mathbf{R}_1 = \text{diag}((\frac{\pi}{8})^2, 5^2, 5^2, 5^2)$, $\mathbf{R}_2 = \text{diag}((\frac{\pi}{8})^2, 0.1^2, 0.1^2, 0.1^2)$. 后者对角度观测引入较大误差, 而对距离观测则比较精确, 此时似然函数是窄形的^[16], 这将增大高斯假设型滤波器的处理难度.

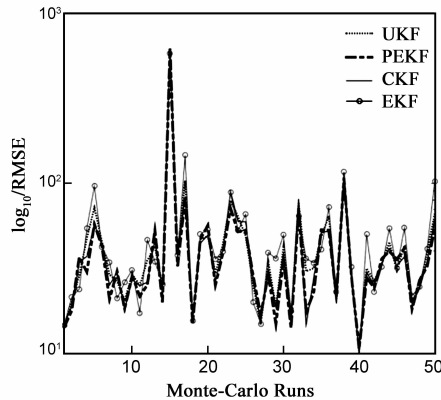


图4 各算法50次MC计算结果($R=R_1$)

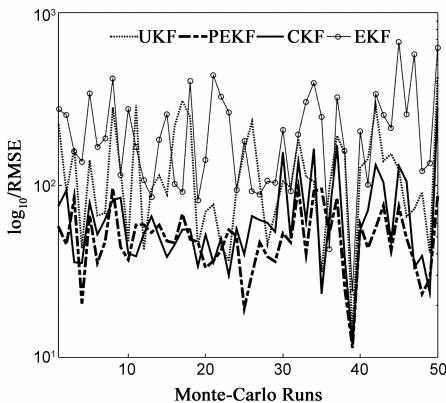


图5 各算法50次MC计算结果($R=R_2$)

表1 RMSE 均值和方差 (300 次 MC 计算)

	$R = R_1$		$R = R_2$	
	均值	方差	均值	方差
EKF	47.07	44.68	188.9	129.8
UKF	41.12	34.62	84.79	72.28
CKF	41.98	44.42	57.98	50.85
PEKF	42.31	42.72	43.83	48.15

图4,5 分别给出了在 $R = R_1$ 和 $R = R_2$ 条件下, PE-

KF 和 EKF, UKF 以及 CKF 经 100 次 MC 仿真得到的 RMSE. 表 1 中给出了两种观测误差条件下 300 次 Monte-Carlo 计算所得的 RMSE 均值和标准差. 由图 4 和表 1 可知在 R_1 条件下, PEKF 性能仍优于 EKF, 且 RMSE 均值接近于 UKF 和 CKF; 而从图 5 中可明显看出, 在 R_2 条件下, 窄状似然函数使得 EKF, UKF 以及 CKF 性能均不同程度下降, 其中 EKF 的性能恶化最为明显. 而 PEKF 在大多数仿真中都给出了最小的 RMSE, 基本保持了其在 R_1 条件下相近的滤波性能, 且误差波动最小. 表明 PEKF 的渐进更新过程能够避免窄状似然函数的对滤波精度和稳定性的影响.

6 总结

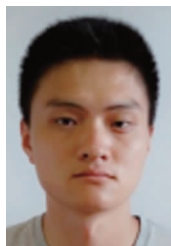
针对窄形似然函数在非线条件下易导致滤波发散的问题, 本文基于渐进贝叶斯框架, 提出渐进扩展卡尔曼滤波器. 通过推导线性高斯条件下的解及相应证明结论, 利用一阶 Taylor 展开导出了非线性条件下滤波器的近似解, 并给出了算法实现, 仿真算例表明, 新滤波器具有较高的精度和合理的时间复杂度. 需要进一步说明的是: 虽然本文提高了 EKF 滤波精度和稳定性, 但非线性条件下渐进贝叶斯解的推导仍基于一阶 Taylor 展开, 且时间更新步骤与 EKF 相同, 因此, 如何在保证状态协方差矩阵误差一致性的前提下提高对于强非线性过程问题的精度是需要继续研究的问题, 若能够推导出 $q(x, \lambda)$ 的时变矩的表达式, 则可直接利用高斯密度积分方法, 以期获得精度更佳的渐进卡尔曼滤波算法. 这将是下一步的工作内容.

参考文献

- [1] Geller D. Linear covariance techniques for orbital rendezvous analysis and autonomous onboard mission planning [J]. AIAA Journal of Guidance, Control and Dynamics, 2006, 29(6): 1404 - 1414.
- [2] Steinbring J, Hanebeck U. Progressive Gaussian filtering using explicit likelihoods [A]. IEEE International Conference on Information Fusion [C]. Salamanca: IEEE, 2014. 1 - 8.
- [3] Perea L, How J, Breger L, et al. Nonlinearity in sensor fusion: Divergence issues in EKF, modified truncated SOF, and UKF [A]. AIAA Conference of Guidance, Navigation, and Control [C]. South Carolina: IEEE, 2007. 3489 - 3499.
- [4] Simon J, Jeffery U, Hugh D. A new method for the nonlinear transformation of means and covariances in filters and estimators [J]. IEEE Trans on Auto control, 2000, 45(3): 477 - 481.
- [5] Arasaratnam I, Haykin S. Cubature Kalman filters [J]. IEEE Trans on Auto Control, 2009, 56(6): 1254 - 1269.
- [6] Li X R, Jilkov V P. A survey of maneuvering target track-

- ing, part VIc: approximate nonlinear density filtering in discrete time [J]. Proceedings of SPIE-The International Society for Optical Engineering, 2012, 8393(1): 2398 – 2415.
- [7] Synder C, Bengtsson T, Bickel P, et al. Obstacles to high-dimensional particle filtering [J]. Monthly Weather Review, 2008, 136(1): 4629 – 4640.
- [8] Daum F, Huang J. Nonlinear filters with particle flow induced by log-homotopy [A]. Proceedings of SPIE, Signal and Data Processing [C]. San Francisco, 2009. 9816 – 9825.
- [9] Daum F. Coulomb's law particle flow for nonlinear filter [A]. Proceedings of SPIE on Signal Processing and Sensor Fusion [C]. San Diego, 2011. 3351 – 3362.
- [10] 张宏欣, 周穗华, 冯士民. 基于弱形式解的粒子流滤波器 [J]. 控制与决策, 2015, 30(5): 853 – 858.
ZHANG Hong-xin, ZHOU sui-hua, FENG Shi-min. A weak solution based particle flow filter [J]. Control & Decision, 2015, 30(5): 853 – 858. (in Chinese)
- [11] Daum F, Huang J. Exact particle flow for nonlinear filters: Seventeen dubious solutions to a first order linear underdetermined PDE [J]. Circuits Systems & Computers, 2010, 45(2): 64 – 71.
- [12] Ding T, Coates M J. Implementation of the Daum-Huang exact-flow particle filter [A]. Statistical Signal Processing Workshop (SSP) [C]. San Diego: IEEE, 2012. 257 – 260.
- [13] Daum F, Huang J. Particle flow with non-zero diffusion for nonlinear filters [A]. Proceedings of SPIE on Signal Processing, Sensor Fusion, and Target Recognition [C]. SPIE, 2013. 4335 – 4341.
- [14] Bar-Shalom Y, XiaoRong Li. Estimation with Applications to Tracking and Navigation [M]. New York: John Wiley&Sons, 2005. 50 – 54.
- [15] Simon D. Optimal State Estimation [M]. New York: John Wiley&Sons, 2006. 22 – 27.
- [16] Bunch P, Godsill S. Approximations of the optimal importance density using Gaussian particle flow importance sampling [OL]. <http://arxiv.org/abs/1406.3183v3>, 2014-11-27.

作者简介



张宏欣 男, 1987 年 12 月出生, 陕西汉中
人. 2010 年毕业于西安理工大学, 现为海军工程
大学博士生, 从事统计信号处理及目标跟踪相
关研究.

E-mail: mylifeforthebattle@hotmail.com



周穗华 男, 1962 年 10 月出生, 广东五华
人, 1984 年毕业于海军工程学院, 1990 年在海军
工程学院获得博士学位. 现为海军工程大学教
授, 从事军用目标特性信息处理及武器系统总
体设计方面研究.