

文章编号:1671-6833(2010)02-0092-05

一种基于粒子滤波的非线性系统参数和状态联合估计方法

王忠勇, 冯卫娜

(郑州大学 信息工程学院, 河南 郑州 450001)

摘要: 提出了一种新的基于粒子滤波的非线性系统参数和状态联合估计方法. 该算法利用粒子滤波方法, 结合核平滑收缩技术, 同时采用标准贝塔分布代替传统的高斯分布, 来拟合系统未知参数的后验分布, 最终实现非线性系统中参数的迭代估计. 仿真结果表明, 该算法提高了未知参数和状态的估计精度, 在估计的收敛性方面也有明显的改善.

关键词: 粒子滤波; 非线性系统; 核平滑收缩; 贝塔分布

中图分类号: TN911.7 **文献标识码:** A

0 引言

非线性非高斯模型的估计问题一直是非线性领域研究的热点. 一个普通的系统估计任务可描述为: 根据系统的观测数据和先验估计信息, 通过一定的算法, 估计出所需要的待估量, 当系统参数未知时, 需同时估计系统的状态和参数. 而扩展 Kalman 滤波状态和参数联合估计算法 JEKF 正是解决此类问题的常用方法^[1-3]. 对于线性高斯系统, JEKF 估计效果是比较好的, 然而, 在处理更为普遍的非线性非高斯对象时, 对于非线性比较强, 且同时含有系统噪声和观测噪声的模型, 估计结果往往会产生较大偏差, 致使估计精度降低, 甚至会产生发散.

粒子滤波方法^[4]是近年来兴起的一种新的估计方法. 粒子滤波的关键思想是通过寻找一组在状态空间中传播的随机样本集, 对后验概率密度函数进行近似, 以样本的均值代替积分运算得到样本的估计及样本的权值. 文献[5]运用两个粒子滤波器对非线性系统参数和状态进行双重估计, 虽然在估计精度方面有了进一步的提高, 但是在运算量和运行时间上也显露出了不足. 笔者在粒子滤波方法的基础上, 通过核平滑收缩方法对参数进行迭代估计, 保留了信息的完整性, 同时避免了参数粒子的退化和贫乏现象. 另外笔者用标准贝塔分布代替了传统提高了参数粒子的采样效

率, 进而提高了参数的收敛速度和估计精度, 而后由动态空间模型得到状态的估计, 最终实现了状态和参数的联合估计. 仿真结果表明, 本文算法能有效地实现非线性系统中状态和参数的联合估计.

1 动态空间模型

在处理很多实际的问题时, 往往需要在知道一些观测数据的情况下, 估计某些未知状态. 假设系统存在未知参数 θ , 且参数为非时变, 需要根据观测量 y_k 同时估计系统状态 x_k 和系统参数 θ . 假设动态系统的状态空间模型为

状态转移模型:

$$x_k = f(x_{k-1}, \theta) + w_k \quad (1)$$

状态观测模型:

$$y_k = h(x_k) + v_k \quad (2)$$

式中: x_k 表示系统在 k 时刻的状态; y_k 表示 k 时刻的观测向量; θ 为系统未知参数; $f(\cdot)$ 和 $h(\cdot)$ 分别表示系统的状态转移函数和测量函数; w_k 和 v_k 为独立的噪声, 分别表示系统的状态噪声和观测噪声. 为完全描述该模型, 假设 k 时刻求得的参数 $\theta = \theta_k$, 通常给出状态的先验概率 $p(x_0)$ 和参数的先验概率 $p(\theta_0)$.

笔者研究的问题是 nonlinear 系统中的联合估计问题, 即在状态空间模型下, 由每一时刻的观测值 y_k 估计联合状态 $z_k = \{x_k, \theta_k\}$. 由贝叶斯理论可知, 联合后验分布 $p(z_{0:k} | y_{1:k})$ 包含了联合状态 $z_{0:k}$

收稿日期: 2009-10-01; 修订日期: 2009-12-29

基金项目: 河南省教育厅自然科学基金项目(2007510019)

作者简介: 王忠勇(1965-), 男, 江西遂川人, 郑州大学教授, 博士, 博士生导师, 主要研究方向: 信息处理、控制理论与应用, E-mail: zywang@zzu.edu.cn.

$= \{x_{0,k}, \theta_{0,k}\}$ 所有的信息,所以,问题就转化为通过观测量 $y_{1:k}$ 估计联合后验概率密度函数 $p(z_{0,k} | y_{1:k})$ [6].

2 粒子滤波及联合估计算法

2.1 粒子滤波原理

粒子滤波的优越性在于它更适合于一般的非线性非高斯系统问题的处理 [7], 该算法的主要思想是使用一个带有权值的粒子集合来表示系统的后验概率密度,然后用这一近似的概率密度表示来计算系统的状态估计. 针对所需要解决的问题,需要用 $\{z_{0,k}^i, \omega_k^i | i=1, \dots, N\}$ 完全描述后验概率密度分布 $p(z_{0,k} | y_{1:k})$, 其中 $\{z_{0,k}^i, i=1, \dots, N\}$ 是支持样本集,相应的权值为 $\omega_k^i, i=1, \dots, N$, 且满足 $\sum_{i=1}^N \omega_k^i = 1$, 而 $z_{0,k} = \{z_j, j=0, \dots, k\}$ 表示到 k 时刻系统所有联合状态的集合.

直接从后验概率密度函数 $p(z_{0,k} | y_{1:k})$ 中采样是比较困难的. 这里引入重要性函数 $q(z_{0,k} | y_{1:k})$ 为概率分布, 与 $p(z_{0,k} | y_{1:k})$ 相近且容易从中采样的分布函数.

为了充分利用 k 时刻以前的估计量,实现对后验分布的递推估计,可以将选择的重要性函数因式分解为:

$$q(z_{0,k} | y_{1:k}) = q(z_k | z_{0:k-1}, y_{1:k}) q(z_{0:k-1} | y_{1:k-1}) \quad (3)$$

则权值递推公式为:

$$\begin{aligned} \omega_k^i &= \frac{p(y_{1,k} | z_k^i) p(z_k^i)}{q(z_k^i | z_{0:k-1}^i, y_{1,k}) q(z_{0:k-1}^i | y_{1,k})} \\ &= \omega_{k-1}^i \frac{p(y_k | z_k^i) p(z_k^i | z_{k-1}^i)}{q(z_k^i | z_{0:k-1}^i, y_{1,k})} \end{aligned} \quad (4)$$

联合后验概率密度函数分布可以用如下经验概率密度分布来近似表示:

$$p(z_{0,k} | y_{1:k}) \approx \sum_{i=1}^N \omega_k^i \delta(z_{0,k} - z_{0,k}^i) \quad (5)$$

当重要性函数 $q(z_k | z_{0:k-1}, y_{1,k}) = p(z_k | z_{k-1}, y_k)$ 时,为最优重要性函数 [8-9], 此时重要性权重为:

$$\omega_k^i = \omega_{k-1}^i \frac{p(y_k | z_k^i) p(z_k^i | z_{k-1}^i)}{p(z_k^i | z_{k-1}^i, y_k)} = \omega_{k-1}^i p(y_k | z_k^i) \quad (6)$$

但是选择最优重要性函数时, $p(y_k | z_{k-1}^i)$ 一般很难求解,因此,需要构造最优重要性函数的近似,选取先验概率密度函数为重要性函数:

$$q(z_k | z_{k-1}, y_{1,k}) = p(z_k | z_{k-1}) \quad (7)$$

此时重要性权重为:

$$\omega_k^i = \omega_{k-1}^i p(y_k | z_k^i) \quad (8)$$

其中,从 $p(z_k | z_{k-1})$ 中随机抽取 N 个有限样本时,因为 $z_k = \{x_k, \theta_k\}$ 为联合状态,所以需要同时实现 x_k 和 θ_k 的粒子更新. 其中笔者采用核平滑收缩方法 [10-11] 和贝塔分布 [12] 采样来实现参数粒子的更新,在下小节中将详细介绍,而后根据式(1)、(2),更新得到状态粒子 x_k^i .

2.2 参数核平滑和贝塔分布采样

利用粒子滤波估计系统未知参数,特别是非时变参数时,由于过于精确的参数预报模型不能增加参数粒子的多样性,普遍存在的问题是经过若干步迭代后,导致参数粒子出现严重的退化和贫乏现象,从而影响状态估计的精度. 针对此类问题,最主流的方法是引入人工噪声来增加粒子多样性,但是增加人工噪声很有可能破坏真实参数后验分布,最终导致粒子集合不再渐进逼近联合后验概率密度函数,所以如何避免信息丢失是解决参数粒子退化和贫乏的最关键的问题 [13].

为了降低人工噪声的干扰作用,保留信息的完整性,在这里采用核平滑收缩技术 [10-11] 来解决非时变参数的迭代估计问题. 用下式来代替真实的后验分布:

$$\rho(\theta_k | Y_{0:k}) \approx \sum_{i=1}^N \omega_k^i \kappa(\theta | m_k^i, h^2 P_k^o) \quad (9)$$

其中, $h > 0$ 表示平滑参数,本文选取 $h = 0.1$; ω_k^i 是 k 时刻第 i 个粒子的重要性权重; $\kappa(\cdot) = \kappa(\theta | m_k^i, h^2 P_k^o)$ 是均值为 m_k^i , 方差为 $h^2 P_k^o$ 的核分布. 核分布有多种选择方法,传统的选择方法为高斯分布,但考虑到高斯分布有比较长的拖尾,采样效率不高,笔者采用标准贝塔分布进行对参数粒子的采样更新.

假设随机变量 Y 服从参数为 (α, β) 的标准贝塔分布 [12], 其概率密度函数为:

$$f(y; \alpha, \beta) = \begin{cases} \frac{1}{B(\alpha, \beta)} y^{\alpha-1} (1-y)^{\beta-1}, & 0 < y < 1 \\ 0, & \text{其他} \end{cases} \quad (10)$$

其中, $B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx$ 是贝塔函数.

当 $\alpha = \beta = 1$ 时,标准贝塔分布为 $[0, 1]$ 上的均匀分布. 标准贝塔分布密度函数的形状可由参数 α, β 来控制,其形状从均匀分布到近似正态分布,从对称到不对称,并且分布在有限的 $[0, 1]$ 范围内. 通过更新参数 α, β , 调整标准贝塔分布在 $[0, 1]$ 区域上的聚集区域和在某一区域的聚集程度,从

而实现对标准贝塔分布函数的控制. 标准贝塔分布和高斯分布比较起来, 标准贝塔分布更能够充分利用其先验信息, 避免高斯分布在拖尾处的采样, 因此具有更高的采样效率.

为了保证在 k 时刻的抽样点和 $k - 1$ 时刻的抽样点具有相同的均值和方差, West 提出一个收缩方法来保证信息的完整性, 即令

$$m_{k-1}^i = a\theta_{k-1}^i + (1 - a)\hat{\theta}_{k-1}, \quad i = 1, \dots, N \tag{11}$$

式中: $\hat{\theta}_{k-1}$ 为 $k - 1$ 时刻参数的加权平均值. 设 P_{k-1}^o 为 $k - 1$ 时刻参数的协方差, $a = \sqrt{1 - h^2}$, $k - 1$ 时刻参数的加权平均值和协方差计算如下:

$$\hat{\theta}_{k-1} = \sum_{i=1}^N \theta_{k-1}^i \omega_{k-1}^i \tag{12}$$

$$P_{k-1}^o = \sum_{i=1}^N (\theta_{k-1}^i - \hat{\theta}_{k-1})^2 \omega_{k-1}^i \tag{13}$$

则 $k - 1$ 时刻第 i 个参数粒子 θ_{k-1}^i 的均值为 m_{k-1}^i , 方差为 $h^2 P_{k-1}^o$, 如果参数 $\theta \in [\theta_m, \theta_n]$, 需把它变换到 $[0, 1]$ 区间^[5], 设 $\phi = \frac{\theta - \theta_m}{\theta_n - \theta_m}$, 则 $\phi \in [0, 1]$, 假设 ϕ 服从标准贝塔分布, 则 $k - 1$ 时刻第 i 个参数粒子 θ_{k-1}^i 所对应的 $[0, 1]$ 区间的 ϕ_{k-1}^i 的均值 μ_{k-1}^i 和方差 σ_{k-1}^2 为

$$\mu_{k-1}^i = \frac{m_{k-1}^i - \theta_m}{\theta_n - \theta_m} \tag{14}$$

$$\sigma_{k-1}^2 = \frac{h^2 P_{k-1}^o}{(\theta_n - \theta_m)^2} \tag{15}$$

则根据 $k - 1$ 时刻 ϕ_{k-1}^i 的均值和方差就可以计算出这一时刻标准贝塔分布的参数 α_{k-1}^i 和 β_{k-1}^i . 表达如下:

$$\alpha_{k-1}^i = \frac{\mu_{k-1}^2 (1 - \mu_{k-1}^i)}{\sigma_{k-1}^2} - \mu_{k-1}^i \tag{16}$$

$$\beta_{k-1}^i = \left[\frac{\mu_{k-1}^i (1 - \mu_{k-1}^i)}{\sigma_{k-1}^2} \right] (1 - \mu_{k-1}^i) \tag{17}$$

则下一个时刻的参数粒子 θ_k^i 所对应的 ϕ_k^i 就可以从贝塔分布 $f(\phi; \alpha_{k-1}^i, \beta_{k-1}^i)$ 中采样抽取得到, 计算 $\theta_k^i = \phi_k^i (\theta_n - \theta_m) + \theta_m$, 而后由 k 时刻的参数粒子 θ_k^i 和 $k - 1$ 时刻的状态粒子 x_{k-1}^i , 根据式(1)和(2)得到 k 时刻的状态粒子 x_k^i , 然后根据 k 时刻的观测量 y_k 计算各个粒子的重要性权重, 就能够得到 k 时刻第 i 个参数粒子 θ_k^i 的均值和方差, 进而更新贝塔分布的参数得到 α_k^i, β_k^i ^[5].

2.3 联合估计算法

通过前面对粒子滤波算法、核平滑收缩及贝

塔分布的介绍, 我们在研究参数和状态联合分布的问题中, 针对参数估计, 为了保证信息的完整性, 使参数尽快的收敛到正确的值, 可以采用核平滑收缩技术, 同时核分布采用标准贝塔分布代替传统的高斯分布, 提高了参数粒子的采样效率. 针对非线性系统参数和状态联合估计改进的算法如下:

设系统状态先验分布为 $p(x_0)$, 参数分布区域为 $[\theta_m, \theta_n]$, 参数初始分布 $p(\theta_0)$ 为均匀分布, 即 $\alpha_0 = 1, \beta_0 = 1$, 设 $h = 0.1$, 初始重要性权重 $\omega_0^i = 1/N, i = 1, \dots, N$.

(1) 初始化: 由初始状态分布 $p(x_0)$ 和初始参数分布 $p(\theta_0)$ 分别抽样得到 N 个状态粒子点 x_k^i 和 N 个参数粒子点 $\theta_k^i, i = 1, \dots, N, 1, 2, \dots$

(2) 更新参数粒子: 由式(11) ~ (17), 计算 α_{k-1}^i 和 β_{k-1}^i , 从贝塔分布 $f(\phi; \alpha_{k-1}^i, \beta_{k-1}^i)$ 抽样得到 ϕ_k^i , 而后计算参数粒子 $\theta_k^i, \theta_k^i = \phi_k^i (\theta_n - \theta_m) + \theta_m$.

(3) 更新状态粒子: 由式(1)和(2)抽样得到 x_k^i .

(4) 更新重要性权重: 根据式(8)计算每一个粒子的重要性权重.

$$(5) \text{ 归一化重要性权重: } \tilde{\omega}_k^i = \omega_k^i / \sum_{j=1}^N \omega_k^j.$$

(6) 状态粒子重抽样: 消除重要性权重较小的状态粒子, 复制权重较大的粒子, 并令每个重抽样之后的状态粒子赋以相同的权重.

(7) 估计: 估计状态值采用 $\hat{x}_k = \sum_{i=1}^N x_k^i \tilde{\omega}_k^i$; 估

计参数值采用 $\hat{\theta}_k = \sum_{i=1}^N \theta_k^i \tilde{\omega}_k^i$.

3 仿真与讨论

为了进一步了解本文算法的估计性能, 采用非线性程度较高且含有系统噪声和观测噪声的状态空间模型, 分别用本文的算法和 JKF 两种方法对状态和参数进行联合估计.

3.1 仿真模型

假设非线性系统的状态空间模型是:

状态转移模型:

$$x_k = 1 + \sin[0.04\pi(k - 1)] + \theta x_{k-1} + w_{k-1}.$$

状态观测模型:

$$y_k = \begin{cases} 0.2x_k^2 + v_k, & k \leq 30 \\ 0.5x_k - 2 + v_k, & k > 30 \end{cases}$$

其中系统噪声 w_k 服从 gamma 分布 $w_k \sim \zeta_0(3, 2)$, 测量噪声 v_k 服从高斯分布, $v_k \sim N(0,$

0.001). 设 $\theta \in [0, 1]$, 试验中样本(粒子)数量为 200, 观测时间为 $t = 60$, 进行 100 次独立实验, 定义单次实验的均方误差为:

$$RMSE = \left[\frac{1}{t} \sum_{i=1}^t (\hat{x}_i - x_i)^2 \right]^{\frac{1}{2}}$$

3.2 仿真结果及分析

图 1 给出了两种算法进行一次独立实验所产生的状态估计结果, 点代表状态真实值, 虚线代表本文算法, 实线代表 JEFK 算法. 可以看出, 本文提出的算法所估计的状态能较好的与真实状态相吻合, 状态估计精度较 JEFK 算法有了进一步的提高. 图 2 和图 3 分别给出了两种算法进行一次独立实验产生的参数估计结果. 为了更明显地看到两种算法的优越性, 本次参数估计比较采用 $t = 2000$. 可以看出, 本文笔者所提出的算法使得参数能够更加准确的收敛到真实值 0.5, 而 JEFK 却不能收敛, 估计精度远远小于本文算法的参数估计精度.

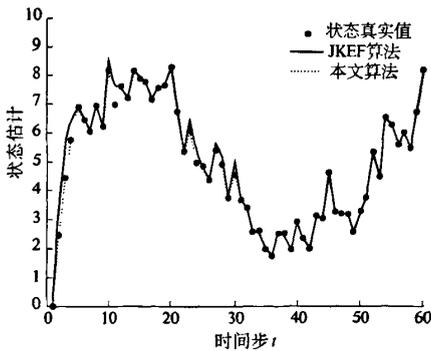


图 1 状态估计结果比较

Fig. 1 Comparison of state estimation

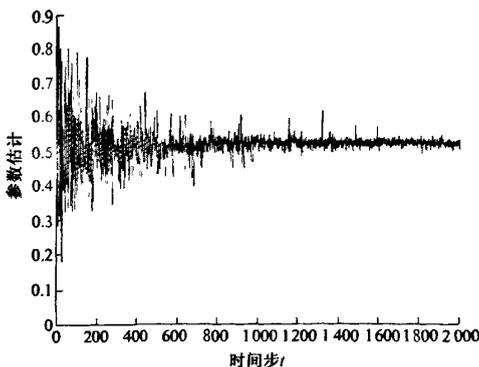


图 2 本文算法参数估计结果

Fig. 2 Parameter estimation of this paper's algorithm

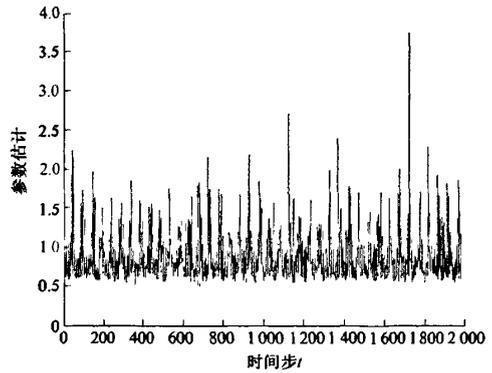


图 3 JEFK 参数估计结果

Fig. 3 Parameter estimation of JEFK

图 4, 5 分别给出了 JEFK 算法和本文算法进行 100 次独立仿真实验后非线性系统状态和参数估计的均方误差比较结果, 表 1 给出了两种算法经过 100 次实验后状态和参数估计的均方误差的均值和方差. 可以看出, 本文提出的算法对状态估计和参数估计的精度及收敛性方面都优于 JEFK 算法.

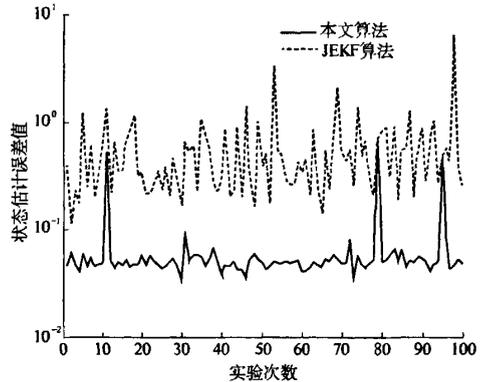


图 4 状态估计均方误差统计

Fig. 4 Mean square error statistics of state estimation

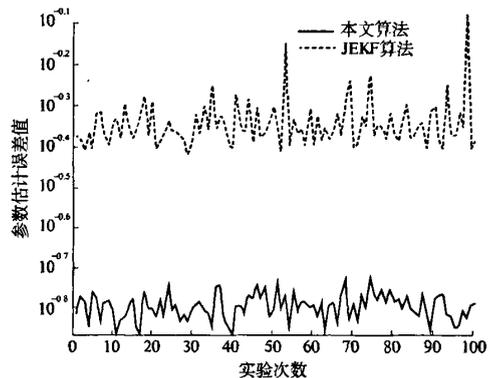


图 5 参数估计均方误差统计

Fig. 5 Mean square error statistics of parameter estimation

表1 两种算法产生的状态和参数的均方误差的均值和方差

Tab.1 The mean square errors' mean and variance of state and parameter produced by two algorithms

估计	算法	RMSE 均值	RMSE 方差
状态估计	JEKF 算法	0.602 790	0.576 630
	本文算法	0.065 889	0.008 012
参数估计	JEKF 算法	0.450 13	0.004 412 1
	本文算法	0.160 32	0.000 137 42

4 结论

采用粒子滤波算法,结合核平滑收缩技术对非时变参数进行迭代估计,并提出用贝塔分布代替传统的高斯分布对参数粒子进行采样,而后由动态空间模型得到状态粒子,实现非线性系统状态和参数的联合估计.实验结果表明,本文提出的算法在估计精度和收敛性方面优于 JEKF 算法.

参考文献:

- [1] WAN E A, NELSON A T. Dual extended Kalman filter methods in kalman filtering and neural networks [M]. New York: John Wiley and Sons Inc, 2001: 123 - 173.
- [2] HABIL S. Combined variable structure and Kalman filtering approach [M]. New York: American Control Conference, 2008: 1855 - 1862.
- [3] 范文兵, 陈达. 卡尔曼滤波器在状态和参数估计中的应用 [J]. 郑州大学学报: 理学版, 2002, 34(4): 44 - 47.
- [4] CARPENTER J, CLIFFORD P, FEARNHEAD P. Improved particle filter for nonlinear problems [J]. IEE Proc Radar Sonar Navig, 1999; 146(1): 1 - 7.
- [5] 侯代文, 殷福亮. 非线性系统中状态和参数联合估计的双重粒子滤波方法 [J]. 电子与信息学报, 2008, 30(9): 2128 - 2133.
- [6] STORVIK G. Particle filters for state - space models with the presence of unknown static parameters [J]. IEEE Transactions on Signal Processing, 2002, 50(2): 281 - 289.
- [7] LIU J S, CHEN R. Sequential Monte Carlo methods for dynamical systems [J]. J of the American Statistical Association, 1998, 93(5): 1032 - 1044.
- [8] DOUCCT A, CODSILL S, ANDRIEU C. On sequential monte carlo sampling methods for Bayesian filtering. Statist [J]. Comput, 2000(10): 197 - 208.
- [9] DE FREITAS J F G, NORANJAN M, GEE A H, et al. Sequential monte carlo methods to train neural network models [J]. Neural Computation, 2000(12): 955 - 933.
- [10] WEST M. Approximating posterior distribution by Mixtures [J]. Journal of Royal Statistical Society, 1993, 55: 409 - 422.
- [11] WEST M. Mixture models, monte carlo, bayesian updating and dynamic models [A]. Proceedings of the 24 the Symposium on the Interface [C]. Fairfax Station, Virginia, 1993: 325 - 333.
- [12] 帕普里斯 A, 佩莱 S. 概率、随机变量与随机过程 [M]. 保铮, 译. 西安: 西安交通大学出版社, 2004: 70 - 72.
- [13] 黄琳. 非线性滤波理论在航天器姿态确定中的应用 [D]. 哈尔滨工业大学 信息工程学院, 2007.

Combined Estimation Method of Parameter and State for Nonlinear Systems Based on Particle Filter

WANG Zhong - yong, FENG Wei - na

(School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: A new combined estimation method of parameter and state for nonlinear systems based on particle filter is proposed. The algorithm uses particle filter methods, combined with the kernel smoothing contraction method, replaces the traditional use of the Gaussian distribution with the standard beta distribution to fit the posteriori distribution of the unknown parameter of the system, in order to achieve the iteration of the parameter estimation of the nonlinear system. Simulation experiment results show that the algorithm improves the estimation accuracy of the the state and the unknown parameter, and the convergence of the estimation has also been significantly improved.

Key words: particle filtering; nonlinear system; kernel smoothing contraction; beta distribution