

## 时变有色观测噪声下基于变分贝叶斯学习的自适应卡尔曼滤波

徐定杰 沈忱\* 沈锋

(哈尔滨工程大学自动化学院 哈尔滨 150001)

**摘要:** 针对卡尔曼滤波中观测噪声是有色的且随时间变化这一情形, 该文提出基于变分贝叶斯学习的自适应卡尔曼滤波算法。该算法先利用差分法, 将时变噪声模型当中的有色观测噪声进行白化处理, 从而使模型转换成了过程噪声与观测噪声相关的白噪声模型。考虑噪声相关条件下的卡尔曼滤波, 并使之与变分贝叶斯学习结合, 将白噪声方差与系统状态变量一起作为参数进行联合的递推估计。仿真结果表明, 该自适应算法对时变的噪声具有较好的跟踪效果, 相对经典卡尔曼滤波有着较高的滤波精度, 最终得到时变有色观测噪声下的状态估计。

**关键词:** 信号处理; 参数估计; 卡尔曼滤波; 自适应滤波; 变分贝叶斯学习

中图分类号: TN911.7

文献标识码: A

文章编号: 1009-5896(2013)07-1593-06

DOI: 10.3724/SP.J.1146.2012.01457

## Adaptive Kalman Filtering with Time-varying Colored Measurement Noise by Variational Bayesian Learning

Xu Ding-jie Shen Chen Shen Feng

(College of Automation, Harbin Engineering University, Harbin 150001, China)

**Abstract:** An adaptive Kalman filtering algorithm based on variational Bayesian learning is suggested to cope with the problem in which colored and time-varying measurement noise is introduced. By use of differencing, the model is converted back to a normal model in which measurement noise is white but correlated with process noise. Kalman filtering is modified owing to the correlation and variational Bayesian learning is combined to jointly estimate the measurement noise and the state in a recursive manner. The simulation results demonstrate that this adaptive algorithm is capable of tracking time-varying noise and provides more accurate state estimation than standard Kalman filtering with colored and time-varying noise.

**Key words:** Signal processing; Parameters estimation; Kalman filtering; Adaptive filtering; Variational Bayesian learning

### 1 引言

经典卡尔曼滤波算法假设噪声是零均值的高斯白噪声, 但是这只适用噪声相关性较弱的情形, 而在大多数诸如导航、定位和制导等工程应用领域当中, 这种相关性并不能够近似忽略, 否则会使滤波效果显著降低。于是, 在系统建模时就有必要考虑有色噪声<sup>[1]</sup>。此外, 卡尔曼滤波器的输入是统计信息确定并保持不变的噪声, 事实上, 噪声统计参数, 或者系统的模型参数, 在实际应用中很有可能是未知的, 且有可能随着系统本身或者量测环境的变化而发生变化。这就需要通过自适应滤波, 在滤波过程中不断更新噪声统计信息或者系统模型结构参数, 以便对感兴趣的系统状态变量进行更好的估计。文献[2]对自适应滤波的方法做了归纳, 其中, 贝叶

斯法可以在理论上解释其它所有方法, 但由于贝叶斯法涉及概率密度函数的计算, 通常由于被积函数形式较为复杂, 或者变量维数较高, 难以得到解析解, 所以贝叶斯法难以在实际应用中得到推广。

正因为直接运用贝叶斯方法对感兴趣参数的概率分布计算较为复杂, Beal<sup>[3]</sup>在他的博士论文中, 在贝叶斯框架下提出了变分贝叶斯学习的概念。它最早源于机器学习领域, 是一种在得到最优解较为困难的前提下, 用近似的方法去寻求次优解。该方法的核心是提议用一个函数形式较为简单的概率分布, 在一定的准则下去不断逼近原分布, 直至两个分布可以近似认为相同。这属于一种确定性近似方法。一方面, 所谓确定性, 就是只要待估计的参数的先验分布属于共轭指数域, 则它的后验分布也一定属于该域, 并拥有着与先验相同的函数形式, 区别在于分布参数不同<sup>[3]</sup>。近似函数正是按照这个原则建立的。它有别于以马尔科夫链蒙特卡洛方法为代

2012-11-12 收到, 2013-02-19 改回

中国博士后科学基金(2012T50330)资助课题

\*通信作者: 沈忱 shen15503@126.com

表的随机性近似方法,通过不断采样获得参数的概率分布<sup>[4]</sup>。此类近似方法的缺点在于采样计算量相对较大,在实时性要求较高的领域内应用十分受限。而同样作为贝叶斯意义下的近似求解方法,变分贝叶斯学习方法收敛速度快,也有着不错的估计精度。另一方面,变分贝叶斯学习假设估计的参数条件独立,对于各变量之间关联性不强的模型,这是一个有效简化计算的方法。譬如,变分贝叶斯学习先后已经用于独立成分分析<sup>[5]</sup>、语音增强<sup>[6]</sup>、盲源分离<sup>[7]</sup>等固定统计模型的参数问题中。近年来,变分贝叶斯学习算法的这些优点逐渐被更多的学者所关注,该方法已经拓展到实时系统中的参数估计与滤波问题中来。Vrettas 等人<sup>[8]</sup>将此方法用于随机系统的参数估计中; Sarkka 等人<sup>[9]</sup>将变分贝叶斯学习用于卡尔曼滤波的白噪声方差估计; Ait-El-Fquih 等人<sup>[10]</sup>将它用于解决卡尔曼滤波中新息矩阵的计算问题;文献[11]在文献[9]的基础上考虑了系统有未知输入的情形,并结合解耦方法,对未知输入、状态变量与噪声方差进行联合估计。受文献[11]启发,考虑到有色噪声在工程应用中更具有一般性,本文以此为假设,对模型进行优化处理,得到噪声相关的卡尔曼滤波模型,通过对该模型的卡尔曼滤波进行调整,再利用变分贝叶斯对系统的模型结构进行学习,实时估计可能变化的白噪声,最终得到有色噪声环境下自适应的变分贝叶斯卡尔曼滤波器。

## 2 变分贝叶斯学习

根据贝叶斯准则,用  $\mathbf{X}$  代表观测数据矢量的集合,  $\Theta$  表示  $d$  个未知模型参数的集合,即  $\Theta = [\theta_1, \theta_2, \dots, \theta_d]$ , 则该参数的后验分布可表示为

$$p(\Theta | \mathbf{X}) = \frac{p(\mathbf{X} | \Theta)p(\Theta)}{p(\mathbf{X})} = \frac{p(\mathbf{X} | \Theta)p(\Theta)}{\int p(\mathbf{X} | \Theta)p(\Theta)d\theta_1 d\theta_2 \dots d\theta_d} \quad (1)$$

如引言所述,当联合概率密度函数  $p(\mathbf{X}, \Theta)$  形式较为复杂,或者  $\Theta$  维数过高时,难以得到后验分布的解析表达式。本文采用的变分贝叶斯学习算法,提议用一个新的分布  $q(\Theta)$  来逼近  $p(\Theta | \mathbf{X})$ 。通常用 Kullback-Leibler (KL) 散度来衡量两个分布之间的差异,即  $\text{KL}(q(\Theta) \| p(\Theta | \mathbf{X}))$ 。而 KL 散度又满足  $\text{KL}(q(\Theta) \| p(\Theta | \mathbf{X})) \geq 0$ , 当且仅当  $q(\Theta) = p(\Theta | \mathbf{X})$  时散度为零。要使  $q(\Theta)$  尽可能逼近真实后验分布  $p(\Theta | \mathbf{X})$ , 就需要最小化  $\text{KL}(q(\Theta) \| p(\Theta | \mathbf{X}))$ 。而 KL 散度满足

$$\text{KL}(q(\Theta) \| p(\Theta | \mathbf{X})) = \int q(\Theta) \ln \frac{q(\Theta)}{p(\Theta | \mathbf{X})} d\Theta \quad (2)$$

对式(2)关于  $\Theta$  求导,得到各个参数的近似分布的通解表达式如式(3)<sup>[12]</sup>:

$$q(\theta_i) = \frac{\exp\left(\mathbb{E}_{q(\theta_{j \neq i})}(\lg p(\mathbf{X}, \Theta))\right)}{\int \exp\left(\mathbb{E}_{q(\theta_{j \neq i})}(\lg p(\mathbf{X}, \Theta))\right) d\theta_j} \propto \exp\left(\mathbb{E}_{q(\theta_{j \neq i})}(\lg p(\mathbf{X}, \Theta))\right) \quad (3)$$

共轭指数域函数的性质可以保证每个计算得到  $q(\theta_i)$  的分布的形式都与其真实后验的形式相一致<sup>[3]</sup>。注意到该通解表达式为隐式解,每个参数  $\theta_i$  的计算需要已知  $\Theta$  内其它参数  $\theta_{j \neq i}$  的概率分布,并在此概率分布下求对数联合概率密度函数关于  $\theta_{j \neq i}$  的期望。在变分贝叶斯学习中,应该首先根据先验知识给超参数赋值,并初始化各个需要估计的参数,然后按式(3)进行迭代计算,直至算法收敛。

## 3 模型概述

在经典卡尔曼滤波适用的状态空间模型中,通常假设观测噪声与过程噪声不相关,并均为零均值白噪声。本文所研究的离散时间的状态空间模型,互相关协方差矩阵将不为零,且观测噪声为有色噪声,其状态方程和观测方程可以表示为

$$\begin{cases} \mathbf{X}_k = \Phi_k \mathbf{X}_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{Z}_k = \mathbf{H}_k \mathbf{X}_k + \mathbf{v}_k \end{cases} \quad (4)$$

其中,过程噪声  $\mathbf{w}_k$  是方差阵为  $\mathbf{Q}_k$  的  $n$  维白噪声序列,观测噪声  $\mathbf{v}_k$  为  $m$  维有色噪声序列。

对于有色噪声,一个基本的白化方法是系统状态增广法,但是这会导致新的观测方程的方差阵为零矩阵,这与本文研究目的矛盾,从而失去自适应滤波的意义。所以,为了使观测噪声的方差阵存在于观测方程,本文采用差分法生成新的白噪声。定义有色噪声序列  $\mathbf{v}_k$  满足式(5):

$$\mathbf{v}_k = \mathbf{V} \mathbf{v}_{k-1} + \mathbf{v}_k^0 \quad (5)$$

其中  $\mathbf{v}_k^0$  为方差阵为  $\mathbf{R}_k$  的白噪声序列,  $\mathbf{V}$  为系数矩阵。再重新定义新的观测  $\mathbf{Z}_k^*$ , 使其满足

$$\mathbf{Z}_{k-1}^* = \mathbf{Z}_k - \mathbf{V} \mathbf{Z}_{k-1} \quad (6)$$

根据状态方程,将等式右边继续展开可以得到

$$\begin{aligned} \mathbf{Z}_{k-1}^* &= \mathbf{Z}_k - \mathbf{V} \mathbf{Z}_{k-1} = (\mathbf{H}_k \Phi_k - \mathbf{V} \mathbf{H}_{k-1}) \mathbf{X}_{k-1} \\ &\quad + \mathbf{H}_k \mathbf{w}_{k-1} + \mathbf{v}_{k-1}^0 = \mathbf{H}_{k-1}^* \mathbf{X}_{k-1} + \mathbf{v}_{k-1}^* \end{aligned} \quad (7)$$

于是得到的新的观测方程  $\mathbf{Z}_k^* = \mathbf{H}_k^* \mathbf{X}_k + \mathbf{v}_k^*$  中的噪声序列  $\mathbf{v}_k^*$  为白噪声序列,新的观测噪声方差阵同样可以通过求期望得到,即

$$\begin{aligned} \mathbb{E}[\mathbf{v}_k^* \mathbf{v}_k^{*\top}] &= \mathbb{E}\left[\left(\mathbf{H}_{k+1} \mathbf{w}_k + \mathbf{v}_k^0\right)\left(\mathbf{H}_{k+1} \mathbf{w}_k + \mathbf{v}_k^0\right)^\top\right] \\ &= \mathbf{H}_{k+1} \mathbf{Q}_k \mathbf{H}_{k+1}^\top + \mathbf{R}_k = \mathbf{R}_k^* \end{aligned} \quad (8)$$

注意到此时  $\mathbb{E}[\mathbf{w}_k \mathbf{v}_k^{*\top}] = \mathbf{Q}_k \mathbf{H}_{k+1}^\top$ , 不为零矩阵,说明

过程噪声与观测噪声相关，需要在后续推导滤波方程时考虑互相关矩阵  $\mathbf{S}_k = \mathbf{Q}_k \mathbf{H}_{k+1}^T$ ，其维数为  $n \times m$ 。

#### 4 改进模型的自适应卡尔曼滤波

当原先含有有色噪声的模型经过白化处理，模型待估计的参数为系统状态  $\mathbf{X}_k$  与新的观测噪声方差阵  $\mathbf{R}_k^*$ 。这样，根据广义贝叶斯滤波理论，自适应滤波问题就转换成了在序贯得到新的观测量  $\mathbf{Z}_k^*$  后，对  $\mathbf{X}_k$  与  $\mathbf{R}_k^*$  的联合递推估计。它们的联合概率密度函数为

$$\begin{aligned} p(\mathbf{X}_k, \mathbf{R}_k^* | \mathbf{Z}_{1:k}^*) &\propto p(\mathbf{X}_k, \mathbf{R}_k^*, \mathbf{Z}_{1:k}^*) \\ &= p(\mathbf{Z}_{1:k}^* | \mathbf{X}_k, \mathbf{R}_k^*) p(\mathbf{X}_k, \mathbf{R}_k^* | \mathbf{Z}_{1:k-1}^*) \end{aligned} \quad (9)$$

而根据 Chapman-Kolmogorov 方程，又有

$$\begin{aligned} p(\mathbf{X}_k, \mathbf{R}_k^* | \mathbf{Z}_{1:k-1}^*) &= \int p(\mathbf{X}_k | \mathbf{X}_{k-1}) p(\mathbf{R}_k^* | \mathbf{R}_{k-1}^*) \\ &\quad \cdot p(\mathbf{X}_{k-1}, \mathbf{R}_{k-1}^* | \mathbf{Z}_{1:k-1}^*) d\mathbf{X}_{k-1} d\mathbf{R}_{k-1}^* \end{aligned} \quad (10)$$

式(9)与式(10)分别是广义贝叶斯滤波下的更新方程与预测方程。一般情况下，先验概率密度函数式(10)中变量的概率分布形式复杂，右侧的积分运算比较困难，从而使后验概率密度函数式(9)也难以得到确定的表达式。但是考虑卡尔曼滤波的适用条件，在贝叶斯意义下，根据先验知识，式(9)中的系统状态、噪声方差以及观测序列的联合概率密度函数  $p(\mathbf{X}_k, \mathbf{R}_k^*, \mathbf{Z}_{1:k}^*)$  可以看作是 3 个独立分布的乘积为

$$\begin{aligned} &p(\mathbf{Z}_{1:k}^* | \mathbf{X}_k, \mathbf{R}_k^*) p(\mathbf{X}_k, \mathbf{R}_k^* | \mathbf{Z}_{1:k-1}^*) \\ &= p(\mathbf{Z}_k^* | \mathbf{X}_k, \mathbf{R}_k^*) p(\mathbf{X}_k | \mathbf{Z}_{1:k-1}^*) p(\mathbf{R}_k^* | \mathbf{Z}_{1:k-1}^*) \\ &= \mathbf{N}(\mathbf{Z}_k^* | \mathbf{H}_k \mathbf{X}_k, \mathbf{R}_k^*) \mathbf{N}(\mathbf{X}_k | \mathbf{m}_k^-, \mathbf{P}_k^-) \\ &\quad \cdot \prod_{i=1}^d \text{IG}(\sigma_{k,i}^2 | \alpha_{k,i}^-, \beta_{k,i}^-) \end{aligned} \quad (11)$$

其中  $\mathbf{N}(\bullet)$  代表高斯分布， $\text{IG}(\bullet)$  代表逆伽马分布， $\sigma_{k,i}^2$  为方差阵  $\mathbf{R}_k^*$  中对角线上第  $i$  个元素。根据共轭指数域函数性质，在变分贝叶斯学习算法中，若用提议分布  $q(\mathbf{X}_k)$  与  $q(\sigma_{k,i}^2)$  分别去逼近  $p(\mathbf{X}_k | \mathbf{Z}_{1:k-1}^*)$  与  $p(\sigma_{k,i}^2 | \mathbf{Z}_{1:k-1}^*)$ ，则它们一定分别服从高斯分布与逆伽马分布。但是考虑到过程噪声与观测噪声相关，式(11)中的分布需要做等价变形，根据文献[13]的分析，可以得到在噪声相关条件下新的观测量的概率分布满足：

$$\begin{aligned} p(\mathbf{Z}_k^* | \mathbf{X}_k, \mathbf{R}_k^*) &= \mathbf{N}\left(\mathbf{Z}_k^* \left| \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right) \mathbf{X}_k, \right. \right. \\ &\quad \left. \left. \mathbf{R}_k^* - \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1} \mathbf{S}_k \right)\right) \end{aligned} \quad (12)$$

其中，上标“-”表示估计值的先验。于是可以运

用通解式(3)的结果分别计算  $q(\mathbf{X}_k)$  与  $q(\mathbf{R}_k^*)$ 。为使计算方便，等号两边取对数。

$$\begin{aligned} &\ln q(\mathbf{X}_k) \\ &= -\frac{1}{2} (\mathbf{X}_k)^T \left[ (\mathbf{P}_k^-)^{-1} + \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right)^T \right. \\ &\quad \cdot \left. \left(\widehat{\mathbf{R}}_k^* - \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1} \mathbf{S}_k\right)^{-1} \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right) \right] \mathbf{X}_k \\ &\quad + (\mathbf{X}_k)^T \left[ (\mathbf{P}_k^-)^{-1} \mathbf{m}_k^- + \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right)^T \right. \\ &\quad \cdot \left. \left(\widehat{\mathbf{R}}_k^* - \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1} \mathbf{S}_k\right)^{-1} \mathbf{Z}_k^* \right] + C \end{aligned} \quad (13)$$

根据式(13)的形式，可以看出  $q(\mathbf{X}_k)$  是一个高斯分布，满足  $q(\mathbf{X}_k) = \mathbf{N}(\mathbf{X}_k | \mathbf{m}_k, \mathbf{P}_k)$ ，其方差阵和均值分别为

$$\begin{aligned} \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{P}_k^- \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right)^T \\ &\quad \cdot \left(\widehat{\mathbf{R}}_k^* - \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1} \mathbf{S}_k + \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right) \right. \\ &\quad \cdot \left. \mathbf{P}_k^- \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right)^T\right)^{-1} \left(\mathbf{H}_k \mathbf{P}_k^- + \mathbf{S}_k^T\right) \end{aligned} \quad (14)$$

$$\begin{aligned} \mathbf{m}_k &= \mathbf{m}_k^- + \mathbf{P}_k^- \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right)^T \\ &\quad \cdot \left(\widehat{\mathbf{R}}_k^* - \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1} \mathbf{S}_k + \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right) \right. \\ &\quad \cdot \left. \mathbf{P}_k^- \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right)^T\right)^{-1} \\ &\quad \cdot \left(\mathbf{Z}_k^* - \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right) \mathbf{m}_k^-\right) \end{aligned} \quad (15)$$

同理有

$$\begin{aligned} \ln q(\mathbf{R}_k^*) &= \sum_{i=1}^d \left[ \left(-\alpha_{k,i}^- - \frac{1}{2} - 1\right) \ln \sigma_{k,i}^2 \right] \\ &\quad - \sum_{i=1}^d \frac{1}{\sigma_{k,i}^2} \left[ \beta_{k,i}^- + \frac{1}{2} \mathbf{E}_{\mathbf{X}_k} \right. \\ &\quad \cdot \left. \left[ \left( \mathbf{Z}_k^* - \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right) \mathbf{X}_k \right)_i \right]^2 \right] + C \end{aligned} \quad (16)$$

所以， $q(\mathbf{R}_k^*)$  是新的  $d$  个相互独立的逆伽马分布的乘积， $q(\mathbf{R}_k^*) = \prod_{i=1}^d \text{IG}(\sigma_{k,i}^2 | \alpha_{k,i}, \beta_{k,i})$ ，其中

$$\alpha_{k,i} = \alpha_{k,i}^- + 1/2 \quad (17)$$

$$\beta_{k,i} = \beta_{k,i}^- + \frac{1}{2} \mathbf{E}_{\mathbf{X}_k} \left[ \left( \mathbf{Z}_k^* - \left(\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}\right) \mathbf{X}_k \right)_i \right]^2 \quad (18)$$

因此，式(15)中的  $\mathbf{R}_k^*$  的估计值可以根据逆伽马分布的性质求得

$$\begin{aligned} \widehat{\mathbf{R}}_k^* &= \text{diag} \left[ \mathbf{E} \sigma_{k,1}^2, \mathbf{E} \sigma_{k,2}^2, \dots, \mathbf{E} \sigma_{k,d}^2 \right] \\ &= \text{diag} \left[ \frac{\beta_{k,1}}{\alpha_{k,1}}, \frac{\beta_{k,2}}{\alpha_{k,2}}, \dots, \frac{\beta_{k,d}}{\alpha_{k,d}} \right] \end{aligned} \quad (19)$$

运用高斯分布的性质  $E[\mathbf{x}_k \mathbf{x}_k^T] = \mathbf{m}_k \mathbf{m}_k^T + \mathbf{P}_k$ ，可以进一步展开式(16)中的期望的表达式

$$E_{\mathbf{X}_k} \left[ \left( \mathbf{Z}_k^* - (\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}) \mathbf{X}_k \right)_i^2 \right] = \left( \mathbf{Z}_k^* - (\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}) \mathbf{m}_k \right)_i^2 + \left( (\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1}) \mathbf{P}_k (\mathbf{H}_k + \mathbf{S}_k^T (\mathbf{P}_k^-)^{-1})^T \right)_{ii} \quad (20)$$

回顾变分贝叶斯参数估计的步骤，只需在每个  $k$  时刻的更新步骤中，利用变分贝叶斯更新方程式(17)与式(18)，对噪声参数进行迭代更新，直至得到最优的估计值。有了更准确的噪声统计信息后，再将该噪声方差的估计值作为输入传递给卡尔曼滤波器，使滤波器能够输出更精确的状态估计。最后进入下一个  $k+1$  时刻的预测步骤。

在预测步骤中，首先考虑系统状态。因为系统状态变量的变化规律满足  $\mathbf{X}_k = \Phi_k \mathbf{X}_{k-1}$ ，而  $\mathbf{X}_k \sim N(\mathbf{X}_k | \mathbf{m}_k^-, \mathbf{P}_k^-)$ ，所以有预测方程

$$\mathbf{m}_k^- = \Phi_k \mathbf{m}_{k-1} \quad (21)$$

$$\mathbf{P}_k^- = \Phi_k \mathbf{P}_{k-1} \Phi_k^T + \mathbf{Q}_k \quad (22)$$

其中  $\mathbf{m}_{k-1}$  与  $\mathbf{P}_{k-1}$  分别为上一时刻通过式(14)与式(15)更新得到的状态的均值与方差阵。

再考虑观测噪声。为了体现观测噪声的变化规律，根据文献[9]的建议，引入一个因子  $\rho \in (0,1)$  来体现该变化，对于观测噪声方差阵  $\mathbf{R}_k^*$  中的  $d$  个对角线元素  $\sigma_{k,1}^2, \sigma_{k,2}^2, \dots, \sigma_{k,d}^2$ ，有

$$\alpha_{k,i}^- = \rho_i \cdot \alpha_{k-1,i} \quad (23)$$

$$\beta_{k,i}^- = \rho_i \cdot \beta_{k-1,i} \quad (24)$$

当  $\rho < 1$  时，方差阵中的每个方差值就变得不稳定，需要进行实时估计。而当  $\rho = 1$  时为特殊情形，噪声为平稳噪声，方差也是平稳的。本文为方便取  $\rho_i = \rho, i = 1, \dots, d$ 。图1显示了算法的具体流程。

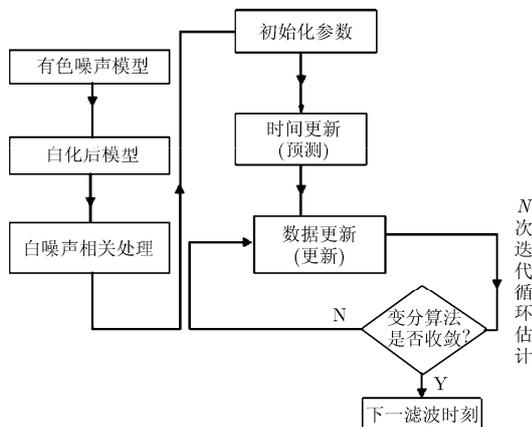


图1 算法整体流程图

### 5 仿真实证

考虑一个随机谐振模型<sup>[9]</sup>，其系统状态方程与观测方程在变化的有色观测噪声假设下如式(25)所示。

$$\begin{cases} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \\ z_k = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + v_k \end{cases} \quad (25)$$

其中，状态变量为3维，并以第1维和第2维的状态作为量测。由于  $v_k$  是有色噪声序列，利用式(5)分离出它的白噪声序列  $v_k^0$ ，使  $v_k = 0.5v_{k-1} + v_k^0$ ；利用式(6)得到  $z_k^* = z_{k+1} - 0.5z_k$ ，取  $z_k^*$  作为新的量测。滤波器状态初始值选为

$$\mathbf{m}_0 = [1 \ 1 \ 1]^T, \mathbf{P}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \alpha_0 = 1, \beta_0 = 1$$

为体现不确定的有色噪声  $v_k$  随时间的变化，模型中使白化的观测噪声的方差不确定，分别在时间段  $[0 \sim 100] \text{ s}$ ,  $[101 \sim 150] \text{ s}$ ,  $[151 \sim 200] \text{ s}$ ,  $[201 \sim 300] \text{ s}$ ,  $[301 \sim 400] \text{ s}$  与  $[401 \sim 500] \text{ s}$  取值为0.1, 0.3, 0.6, 0.2, 1.0与0.3。

图2虚线反映白噪声方差的真实变化，黑色实线为本文的自适应估计曲线，蓝色实线为文献[9]中自适应算法的曲线。从该图中可以看出，两条实线在整个500s时间段内都能够追踪真实虚线，即滤波器可以自适应地跟踪方差，输出随方差变化而变化。注意到在大多数时间里，黑色曲线对方差的跟踪情况要好于蓝色曲线，这是因为实验模拟的有色噪声环境超出了针对白噪声的自适应算法所能起作用的范围，因而其对方差的估计精度不如有色噪声的自适应算法。此外，注意到每次方差的真实值发生跳变后，滤波器自适应的噪声估计值需稍作调整才会跟踪上真实方差，这是由于变分贝叶斯需要通过系统的学习，进行迭代计算之后才会逐渐收敛，而当外界突然发生变化，输入的初始值偏离真实值较多时，就会花费相对较多的计算时间才会收敛。因此，合理的初始值只会对变分贝叶斯学习算法的收敛速度产生影响，但是不影响最终的估计效果。

为检验基于变分贝叶斯学习的自适应卡尔曼滤波器在有色观测噪声下的状态估计效果，本实验还将其与模型白化和相关性处理之后的经典卡尔曼滤波器以及文献[9]中所提出的针对白噪声的自适应卡尔曼器进行比对。图3(a)~图3(c)显示了3个状态变量的滤波效果对比图。蓝色实线代表了状态的真

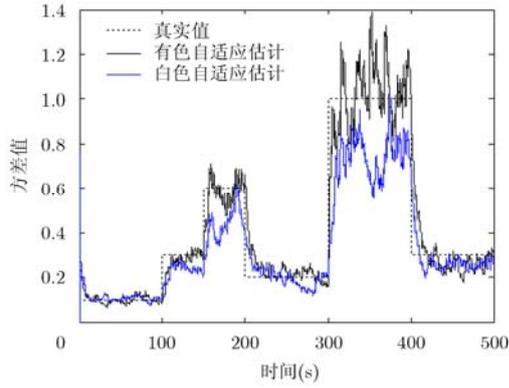


图2 变分贝叶斯学习方差跟踪曲线

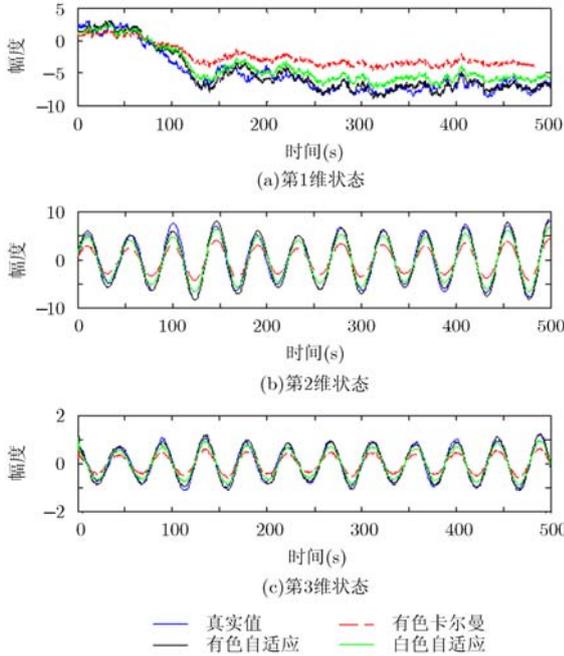


图3 3维状态滤波对比图

实值，黑色实线代表本文的基于变分贝叶斯学习的自适应卡尔曼滤波曲线，红色虚线为传统卡尔曼滤波算法的滤波曲线，绿色实线为针对白噪声的自适应算法曲线。

可以看出，在噪声统计特性未知的情形下，两条自适应的滤波曲线相比传统卡尔曼滤波曲线，在整个仿真时间段内能够更好地对状态进行估计，滤波精度较高。相反，卡尔曼滤波器的曲线则误差较大。从图3的第1维状态估计还可以看出，由

于最初是在第100s时刻改变了观测噪声的方差，卡尔曼滤波器并不能检测其变化，与100s时刻之后的状态跟踪效果相比前100s急剧恶化，之后也并没有减小误差的趋势，基本不能对噪声变化进行跟踪，当后续状态的变化规律发生变化后，也不能对其进行很好的估计。其它两维状态的估计效果表现的也同样不理想，在噪声变化的环境下，自始至终都有着比较大的估计误差。而自适应的曲线无论在哪一个状态，因为能够跟踪变化的方差，对状态估计的效果总体较为理想，误差也相对较小。更进一步，从3个状态曲线图可以看出，在有色噪声环境下，有色噪声的自适应曲线比白噪声的自适应曲线拥有更高的滤波精度，这是因为后者的算法是针对白噪声设计的，而在有色噪声环境下，性能会发生退化。但是后者的滤波精度仍要高于针对有色噪声的卡尔曼滤波。可以得出结论，对于噪声统计的不确定性，其影响要大于把有色噪声当作白噪声来进行自适应处理。

此外，为了量化地体现3种滤波算法的误差对比，实验计算了3种算法在500s时间段内状态的均方根误差(Root Mean Square Error, RMSE)，考虑实验模型，采用的计算公式为

$$RMSE(\mathbf{X}) = \sqrt{\frac{\sum_{k=1}^{500} [(x_{k1} - \hat{x}_{k1})^2 + (x_{k2} - \hat{x}_{k2})^2 + (x_{k3} - \hat{x}_{k3})^2]}{500}}, \quad i = 1, 2, 3 \quad (26)$$

经过计算，得到500s卡尔曼滤波的RMSE为3.8542，而有色噪声自适应滤波的RMSE为1.1071，白色噪声自适应滤波的RMSE为1.6869。为更清楚地显示两种自适应滤波与卡尔曼滤波的误差值，分别再对图3中的3个状态量求RMSE，将所得结果列于表1中。通过定量计算可以看出，对于整个仿真时间而言，针对有色噪声的自适应滤波算法处理得到3个状态估计的误差均远远要小于卡尔曼滤波的估计误差，并且优于针对白色噪声的自适应滤波算法。这一结论也可以从之前的图3中定性分析获得。

## 6 结论

对于有色观测噪声假设且噪声统计特性随时间

表1 仿真时间内各状态变量RMSE对比

时间	第1维有色自适应	第1维白色自适应	第1维有色卡尔曼	第2维有色自适应	第2维白色自适应	第2维有色卡尔曼	第3维有色自适应	第3维白色自适应	第3维有色卡尔曼
[1~500] s	0.8156	1.3310	3.0997	0.7398	1.0234	2.2677	0.1142	0.1639	0.3227

变化的状态空间模型, 本文利用差分法将其转化为白噪声相关条件下的自适应滤波问题, 并结合改进的卡尔曼滤波算法与变分贝叶斯学习对观测噪声与状态变量进行联合的递推估计。仿真实验的结果表明, 相比于卡尔曼滤波, 该算法在有色时变噪声条件下有着更强的稳定性与自适应性, 拥有较高的滤波精度。

### 参 考 文 献

- [1] 付梦印, 邓志红, 闫莉萍. Kalman 滤波理论及其在导航系统中的应用[M]. 第 2 版, 北京: 科学出版社, 2010: 58-61.  
Fu Meng-yin, Deng Zhi-hong, and Yan Li-ping. Kalman Filtering Theory and Its Applications in Navigation System[M]. 2nd Edition, Beijing: Science Press, 2010: 58-61.
- [2] Mehra R. Approaches to adaptive filtering[J]. *IEEE Transactions on Automatic Control*, 1972, 17(5): 693-698.
- [3] Beal M J. Variational algorithms for approximate Bayesian inference[D]. [Ph.D. dissertation], University College London, 2003.
- [4] Andrieu C, Doucet A, and Holenstein R. Particle Markov chain Monte Carlo methods[J]. *Journal of the Royal Statistical Society*, 2010, 72(3): 269-342.
- [5] Lawrence N D and Bishop C M. Variational Bayesian independent component analysis[R]. Technical Report, Computer Laboratory, University of Cambridge, 2000.
- [6] Huang Qing-hua, Yang Jie, and Zhou Yue. Variational Bayesian method for speech enhancement[J]. *Neurocomputing*, 2007, 70(16-18): 3063-3067.
- [7] Sun Shi-jun, Peng Cheng-lin, Hou Wen-sheng, *et al.* Blind source separation with time series variational Bayes expectation maximization algorithm[J]. *Digital Signal Processing*, 2012, 22(1): 17-33.
- [8] Vrettas M D, Cornford D, and Oppor M. Estimating parameters in stochastic systems: a variational Bayesian approach[J]. *Physica D*, 2011, 240(23): 1877-1900.
- [9] Sarkka S and Nummenmaa A. Recursive noise adaptive Kalman filtering by variational Bayesian approximations[J]. *IEEE Transactions on Automatic Control*, 2009, 54(3): 596-600.
- [10] Ait-El-Fquih B and Rodet T. Variational Bayesian Kalman filtering in dynamical tomography[C]. IEEE International Conference on Acoustics, Speech, and Signal Processing, Prague, Czech Republic, 2011: 4004-4007.
- [11] Sun Jun-long, Zhou Jie, and Gu Xue-kui. Variational Bayesian two-stage Kalman filter for systems with unknown inputs[J]. *Procedia Engineering*, 2012, 29: 2265-2273.
- [12] Bishop C M. Pattern Recognition and Machine Learning[M]. Berlin: Springer, 2006: 461-466.
- [13] Simon D. Optimal State Estimation[M]. Hoboken, New Jersey: John Wiley & Sons, 2006: 183-188.

徐定杰: 男, 1966 年生, 教授, 研究方向为无线电及卫星导航技术、雷达抗干扰技术。  
沈 忱: 男, 1988 年生, 博士生, 研究方向为卫星导航技术、统计信号检测与处理。  
沈 锋: 男, 1981 年生, 副教授, 研究方向为信号检测技术、导航抗干扰技术。