Adaptive Kalman Filter for Noise Estimation and Identification with Bayesian Approach

Farhad Asadi, S. Hossein Sadati

Abstract—Bayesian approach can be used for parameter identification and extraction in state space models and its ability for analyzing sequence of data in dynamical system is proved in different literatures. In this paper, adaptive Kalman filter with Bayesian approach for identification of variances in measurement parameter noise is developed. Next, it is applied for estimation of the dynamical state and measurement data in discrete linear dynamical system. This algorithm at each step time estimates noise variance in measurement noise and state of system with Kalman filter. Next, approximation is designed at each step separately and consequently sufficient statistics of the state and noise variances are computed with a fixed-point iteration of an adaptive Kalman filter. Different simulations are applied for showing the influence of noise variance in measurement data on algorithm. Firstly, the effect of noise variance and its distribution on detection and identification performance is simulated in Kalman filter without Bayesian formulation. Then, simulation is applied to adaptive Kalman filter with the ability of noise variance tracking in measurement data. In these simulations, the influence of noise distribution of measurement data in each step is estimated, and true variance of data is obtained by algorithm and is compared in different scenarios. Afterwards, one typical modeling of nonlinear state space model with inducing noise measurement is simulated by this approach. Finally, the performance and the important limitations of this algorithm in these simulations are explained.

Keywords—Adaptive filtering, Bayesian approach Kalman filtering approach, variance tracking.

I. INTRODUCTION

The Kalman Filter (KF) can be estimated dynamical state from noisy measurements. In this method, dynamic and measurement processes can be approximated by linear Gaussian state space models. This model is a practical model in engineering due to the modeling of various noises where Gaussian white noise corrupted the measurements [1]. The extended Kalman filter (EKF) and the unscented Kalman filter (UKF) encompass this method to nonlinear dynamical states and measurement. The EKF employs a Kalman filter for system dynamics that results from the linearization of the original nonlinear filter dynamics around the previous state estimates by forming a Gaussian approximation to the posterior state distribution in the modeling [2]–[4]. A serious constraint in EKF and UKF is that they adopt prior knowledge of the measurement and the parameters of the dynamical model, including the noise statistics status.

In different signal processing applications, there are many sources of interference and noise in systems, and in these conditions, the efficiency of the algorithm for computation and estimation is crucial [5]–[7]. In general, the exact knowledge of the noise statistics characteristics is not obvious, and we don’t have exact information about it in many practical situations. GPS positioning systems or fault detection systems are systems with these properties [8], [9].

For solving this problem and solving uncertain parameters in the model, there is a different algorithm that among those, Adaptive Kalman Filters (AKF) are common in literature [10]. This approach can be estimated noise statistics characteristics or noise variances. In other words, it computed variance-covariance matrices relating to the state and the measurement models. Moreover, the estimation of dynamical states and measurement can be done simultaneously [11]. Intuitively, in AKF, the algorithm adjusts its knowledge about state and measurement matrices values according to the gap between predicted estimates and the current measurements.

In literature, different adaptive filtering approaches are divided to four categories; Bayesian, maximum likelihood, correlation analysis and also covariance matching methods [12]. The first two categories assume the noise covariance estimation problem as a parameter estimation problem.

Bayesian approach is more common in respected to the other approaches and in different computational signal processing, this approach is used. As said before, estimation of uncertainty with dynamical states is important in filtering problems and Bayesian approach is a strong method for approximation of posterior status of these disturbances. In the Bayesian inference approach, the posterior probability density function (pdf) is computed from their noise covariance matrix by applying the Bayes’ formula recursively [13].

Some algorithms like particle filter [14], [15] used Bayesian formulations for noise adaptive filtering. On the other hand, reference like [16] is used and developed Bayesian approach based on approximation of posterior distribution and one of the important advantage of this algorithm is related to low computational cost time. Moreover, references like [17] is developed an approach for recursive Bayesian inference and its approach is suitable for signal processing applications and also for nonlinear dynamical system approach [18], [19].

In recent years, an approximation algorithm for linear and nonlinear state space models with unknown and varying variances is proposed. In references [20], [21] a Kalman smoother with variational structure is proposed for approximation of stationary noises. On the other hand, in reference [22], a fixed form approach for models with time varying variances is proposed. One disadvantage of its

Farhad Asadi is with the K.N. Toosi University of Technology, Iran (phone: +989357430949, e-mail: farhadasadi1990@gmail.com).

S. Hossein Sadati is with the K. N. Toosi University of Technology, Tehran, Iran (e-mail: sadati@kntu.ac.ir).
approach is preparation of exact model of dynamical system and also statistical information should be accurate and available for the algorithm.

In general, the Bayesian approach suffers computationally because of the numerical integration over a very large parameter space. In reference [23], Bayesian adaptive Kalman filter is formulated and it can be used for variances in measurement and with dynamical state. But in this paper, we developed a series of simulation in conditions when the variance of measurement has different distribution and statistical characteristics and then performance of approach for following and tracking this variance is developed. Also, a nonlinear state-space model is applied to these approaches, and estimation of states and measurement of system with adaptive Kalman filter is investigated. In many references, this method is used extensively regarding the approximation of the joint statistical characteristics and then performance of approach for estimation of variance is simulated. Finally, in this paper, we developed a series of simulation in conditions when the variance of measurement has different distribution and statistical information should be accurate and available for the algorithm.

Finally, by having the measurement y_k, the predictive distribution above is updated to a posterior distribution by the ISNI:0000000091950263

III. PROBLEM FORMULATION

Summary formulation of algorithm is explained in the below.

A. Overall modeling of Kalman filter

The discrete-time linear state space model can be considered here as below in (1).

\[ x_k = A_k x_{k-1} + q_k \]
\[ y_k = H_k + r_k \]

where \( q_k \sim N(0, Q_k) \) is the Gaussian process noise, \( r_k \sim N(0, \Sigma_k) \) is the measurement noise with assumed covariance \( \Sigma_k \), and the initial state has a prior Gaussian distribution \( x_0 \sim N(m(0), P(0)) \).

The measurement \( y_k \) is a d-dimensional vector and the \( x_k \) is an n-dimensional vector. Here, \( x_k \) is an unknown variable and \( y_k \) is an observed variable. Time is shown by \( k \) in the matrices of \( A_k \), \( H_k \), and they are assumed to be known and also with \( \Sigma_k \). \( H_k \) is measurement matrix and \( \Sigma_k \) is the measurement noise covariance matrix. Also, the parameters of the initial state \( m_0 \), \( P_0 \) are assumed to be known in initial condition. The estimation of states is recursively applied using two Kalman filter steps as below.  

1. Prediction step:

\[ m_k^- = A_{k-1} m_{k-1} \]
\[ P_k^- = A_{k-1} P_{k-1} A_{k-1}^T + Q_{k-1} \]

2. Update step:

\[ S_k = H_k P_k^- H_k^T + \Sigma_k \]
\[ K_k = P_k^- H_k^T S_k^{-1} \]
\[ C = m_k^- + K_k (y_k - H_k m_k^-) \]
\[ P_k = P_k^- - K_k S_k K_k^T \]

where \( m_k^- \) is the a priori state mean, \( m_k \) is the posteriori state mean, \( P_k^- \) is the priori state covariance and \( P_k \) is the posteriori state covariance. In this algorithm, observation noise variance parameters, \( \sigma^2_{z_k} = 1 \ldots d \), are stochastic with independent dynamic models. We represent the diagonal covariance matrix for this parameter as \( \Sigma_z = \text{diag} \left( \sigma^2_{z_1}, \ldots, \sigma^2_{z_d} \right) \). Also, the construction of a suitable dynamical model of the observation noise variances is represented by \( p (\Sigma_z | \Sigma_{z-1}) \). Dynamic models of the states and the variance parameters are assumed independent according to below equation.

\[ p (x_k, \Sigma_z | x_{k-1}, \Sigma_{z-1}) = p (x_k | x_{k-1}) p (\Sigma_z | \Sigma_{z-1}) \]

The objective of Bayesian optimal filtering of the above model is to calculate the posterior distribution \( p (x_k, \Sigma_z | y_{1:k}) \). Generally, the well-known recursive solution to this filtering problem consists of the following steps [23]. Firstly, the recursion begins from the prior distribution \( p (x_k, \Sigma_z) \). Next, the predictive distribution of the state and measurement noise covariance \( \Sigma_k \) is specified by the Chapman-Kolmogorov equation. Finally, by having the measurement \( y_k \), the predictive distribution above is updated to a posterior distribution by the...
Bayes’ rule that is written in (5).

\[ P (x_k, \Sigma_k \mid y_{1:k}) \propto p (y_k \mid x_k, \Sigma_k) p (x_k, \Sigma_k \mid y_{1:k-1}) \]  

(5)

In [23] the recursion and suitable dynamics for the observation noise variances for the posterior update is proposed.

B. Bayesian Inference Formulation

Conditional distribution for \( x_{k-1} \) and \( \Sigma_{k-1} \) is computed from the measurements \( y_1 \ldots y_{k-1} \). An Independent Inverse-Gamma distribution is modeled as follows in (6).

\[ P \left( x_{k-1}, \Sigma_{k-1} \mid y_{1:k-1} \right) \propto N \left( x_{k-1} \mid m_{k-1}, \Sigma_{k-1} \right) \times \prod_{i=1}^{d} \text{INV-Gamma} \left( \sigma^2_{k,i} \mid \alpha_{k,i-1}, \beta_{k,i-1} \right) \]  

(6)

This approximation is chosen, because the Inverse-Gamma distribution is the prior distribution. Also, using an Inverse-Gamma model for variances of Gaussian models is common in Bayesian analysis [23] because the dynamics of the state and observation noise variances are assumed to be independent. It should be said in the posterior update step, the state and observation noise variance parameters will be coupled through the like-hood distribution \( p (y_k \mid x_k, \Sigma_k) \). We follow the standard variational Bayesian approach for a free form approximate distribution in \( P (x_k, \Sigma_k \mid y_{1:k}) \) as follows:

\[ Q \left( x_k, \Sigma_k \mid y_{1:k} \right) \approx Q_k (x_k) Q_k (\Sigma_k) \]  

(7)

By forming the standard variational Bayesian (VB) approach [23] the finalized prediction updated cycle is obtained as follows.

\[ Q_k (x_k) = N \left( x_k \mid m_k, P_k \right) \]  

\[ Q_k (\Sigma_k) = \prod_{i=1}^{d} \text{INV-Gamma} \left( \sigma^2_{k,i} \mid \alpha_{k,i-1}, \beta_{k,i-1} \right) \]  

(8)

(9)

where the parameters \( m_k, P_k, \alpha_{k,i}, \beta_{k,i} \) are obtained by following equations.

\[ m_k = \bar{m}_k + P_k \Sigma_k^{-1} (y_k - H_k m_k) \]  

\[ P_k = P_k - P_k \Sigma_k^{-1} P_k H_k \Sigma_k^{-1} (y_k - H_k m_k) \]  

\[ \alpha_{k,i} = \alpha_{k,i-1} + 1 \]  

\[ \beta_{k,i} = \beta_{k,i-1} + \frac{1}{2} \left[ (y_k - H_k m_k)^2 + (m_k P_k H_k)^2 \right] \]  

(10)

where \( i=1, \ldots, d \) and

\[ \Sigma_k = \text{diag} \left( \beta_{k,1} / \alpha_{k,1}, \ldots, \beta_{k,d} / \alpha_{k,d} \right) \]  

(11)

Also dynamic model of noise variance usually is not defined but it can be modeled by approximate posteriors. First in algorithm expected measurement noise precisions is considered constant, and then their variances are increased by a factor of \( \rho \) \((\rho \in (0, 1])\). This is obtained by following equations as follows.

\[ \alpha_{\omega,i} = \rho \alpha_{\omega,i-1} \]  

\[ \beta_{\omega,i} = \rho \beta_{\omega,i-1} \]  

(12)

(13)

In these equations \( \rho = 1 \) correspond to stationary variances and lower values increase their assumed fluctuation. In the modeling if the cross correlation between the prediction and observation error is ignored, then covariance becomes diagonal matrix and in many practical situation it is a proper assumption. The fixed point iteration of the algorithm is computed as bellows. Firstly, the prediction of the parameters of the predicted distribution is as follows.

\[ m_k = A_k m_{k-1} \]  

\[ P_k = A_k P_{k-1} A_k^T + Q_k \]  

\[ \alpha_{k,i} = p_1 \alpha_{k-1,i} \]  

\[ \beta_{k,i} = p_1 \beta_{k-1,i} \]  

(14)

Then in the update section firstly we set \( m_k^{(0)} = m_k \) , \( P_k^{(0)} = P_k \) , \( \alpha_{k,i}^{(0)} = \frac{1}{2} + \alpha_{k,i} \) and \( \beta_{k,i}^{(0)} = \frac{1}{2} + \beta_{k,i} \) for \( i=1, \ldots, d \). Next, iterate the following equations in N steps:

\[ \Sigma_k^{(n)} = \text{diag} \left( p_2^{(n)} \alpha_{k,1}^{(n)}, \ldots, p_2^{(n)} \alpha_{k,d}^{(n)} \right) \]  

\[ m_k^{(n+1)} = m_k + P_k \Sigma_k^{-1} (y_k - H_k m_k) \]  

\[ P_k^{(n+1)} = P_k - P_k \Sigma_k^{-1} P_k H_k \Sigma_k^{-1} (y_k - H_k m_k) \]  

\[ \beta_{k,i}^{(n+1)} = \beta_{k,i}^{(n)} + \frac{1}{2} \left[ (y_k - H_k m_k)^2 + (H_k P_k H_k)^2 \right] \]  

(15)

And set \( \beta_{k,i}^{(n)} = \beta_{k,i}^{(n)} \), \( m_k = m_k^{(N)} \) , \( P_k = P_k^{(N)} \) .

In general, the algorithm should be started from a prior of the form:

\[ P \left( x_0, \Sigma_0 \right) = N \left( x_0 \mid m_0, P_0 \right) \prod_{i=1}^{d} \text{Inv-Gamma} \left( \sigma^2_{0,i} \mid \alpha_{0,i}, \beta_{0,i} \right) \]  

(16)

And the approximation formed by the algorithm on step k is as follows.

\[ P \left( x_k, \Sigma_k \mid y_{1:k} \right) \approx N \left( x_k \mid m_k, P_k \right) \prod_{i=1}^{d} \text{Inv-Gamma} \left( \sigma^2_{k,i} \mid \alpha_{k,i}^{(n)}, \beta_{k,i}^{(n)} \right) \]  

(17)

IV. Experimental Results

In this section, different simulation results are explained. Firstly, conditions when noise variance cannot be estimated by approach are presented in two different situations. Then, simulation results is done for adaptive Kalman filter with ability of approximation of variance in measurement of data and influence of variance in measurement and its distribution is discussed. Also, in simulation results the variance of measurement is increased to show the effect of variance in the algorithm. This artificial data has time varying error and also has unknown time varying variance. An example of the time
varying nature of the errors involved, is the initialization of the sensor error states.

A. Adaptive Kalman Filter without Ability of Approximation of Variances

The first simulated data is shown in Fig. 2 and it is related to measurement of system with its estimation and then estimated variance trajectory with true variance is compared and plotted in Fig. 3. Also, the default variation of true variance is as follows. Firstly, in the simulation the measurement noise has the variance of 0.2 and in the time step of 100 the variance quickly is increased to 1.45 and around time 200; it again quickly decrease to value 0.7. Because of wrong initial condition in variance the well matching is not obtained. But, in the adaptive Kalman filter the transition probabilities can be chosen in such a way that probability of switching mode from one mode to another is matched to the variation of variance with some try and error.

In the second simulation, due to the higher variance noise in measurement data, the performance of Kalman filter for variance following is poor and this makes that variational Bayesian for approximating of noise and variance is more important.

B. Adaptive Kalman Filter with Ability of Approximation of Variances with Erupted Initial Condition

In this section adaptive Kalman filter has ability of variance tracking of noise measurement data and it uses the variational Bayesian approach. Also, a Gaussian random variable with unknown time varying variance $\sigma^2$ is applied. Firstly, the Variance trajectory in different initial conditions is developed in measurement data and in this situation again the variance following is erupted with unrelated initial condition. This irregularity in initial condition has important effect on variance noise tracking in measurement, because the resulted error due to this condition cannot be removed in short steps of algorithm. Therefore, it causes deviation from the exact variance trajectory, and these conditions in three different simulations are modeled as bellows.

Estimation of measurement and dynamical states with this approach for three different simulations are plotted in Figs. 6 to 8 and respected trajectory following with adaptive Kalman filter are plotted in Figs. 9 to 11 respectively.
Although this approach uses approximation to the variance noise distribution and forms a Gaussian state distribution conditionality in each time step, but in high-dimensional data with irrelevant variance structure, the assumption of center of limit for modeling of this approach is not practical well. Furthermore, due to the recursive nature of the filter estimation, the performance of the filter is dependent on a priori estimate. This means that the adaptive filter is not entirely self-tuning so we should reconsider the dimension of data for using limitation of this approach according to the above simulations.

C. Second Modeling for Simple Pendulum with Noise Disturbance

In this section, the continuous-discrete sequential is applied to estimation of a partially measured simple pendulum model which is distorted by a random noise term. The stochastic differential equation for the angular position of a simple pendulum, which is distorted by random white noise accelerations $w(t)$ with spectral density $q$ can be written as belows.

$$\frac{d^2x}{dt^2} + a^2 \sin(x) = w(t)$$  \hspace{1cm} (18)

If we define the state as $x = (x, \frac{dx}{dt})$ then, it is changed to state space form and the model can be written as belows.

$$\frac{dx_1}{dt} = x_2$$ \hspace{1cm} (19)

$$\frac{dx_2}{dt} = -a^2 \sin(x_1)dt + d\beta$$ \hspace{1cm} (20)

Assume that the state of the pendulum is measured once per unit time and the measurements are disturbed by Gaussian measurement noise with an unknown variance $\sigma^2$ then a suitable model in this case can be written as belows.

$$y \sim N(x_1(t), \sigma^2)$$

$$\sigma^2 \sim \text{Inv-}X^2(v, \sigma^2)$$  \hspace{1cm} (21)
In Fig. 12 this model with measurement data, actual and estimated signal are plotted. According to the Fig. 12, it is evident that the estimation of signal is most generated in area with high concentration of data and in this area the correlation of our data is higher so this algorithm can detect this important information for tracking and estimation of our signal. In this situation when high level of noise is inputted in measurement of data, this algorithm cannot follow the true signal well and this simulation is plotted in Fig. 13. In summary, when this method is chosen in high dimensional data with high noise variance of measurement, limitation of this approach to this condition should be investigated properly in the algorithms.

![Fig. 12 Distribution of measurement with estimation of signal](image1)

![Fig. 13 Distribution of measurement with estimation of signal with more noise variance](image2)

V. CONCLUSION

In this article, we have presented adaptive Kalman filtering algorithm, which is based on recursively forming approximations to the joint distribution of state and noise parameters. The performance of the different variance measurement has been demonstrated in a simulated application. Then, simulation is executed to adaptive Kalman filter with ability of noise variance tracking in measurement data. In these simulations, the effect of noise distribution of measurement data in each step is calculated and true variance of data is obtained by algorithm and is compared in different scenarios. Then, limitation and performance of this approach in high dimensional data are simulated and discussed.

REFERENCES


