Cubature Kalman Filter with Risk Sensitive Cost Function

Shovan Bhaumik #1, and Swati #2

# Electrical Engineering Department
Indian Institute of Technology Patna
Patna, Bihar, India

1 shovan.bhaumik@iitp.ac.in
2 swati_iitp@iitp.ac.in

Abstract—A novel method to optimize risk sensitive cost function for nonlinear system based on cubature quadrature rule has been proposed in this paper. The proposed filter has been named as risk sensitive cubature Kalman filter (RSCKF). Also a simple and easy to follow derivation of cubature quadrature rule for multi dimensional integral has been provided. Although the computational load is comparable with extended risk sensitive filter (ERSF), the proposed filter is able to overcome the inherent disadvantages associated with it. The theory and formulation of proposed RSCKF have been presented in this paper. Due to more accuracy, enhanced robustness and computational efficiency compare to ERSF, the proposed robust estimator may find place for on-board real life applications.

Index Terms—Risk sensitive Kalman filter, Nonlinear estimation, Robust filtering

I. INTRODUCTION

The proposed robust estimator minimizes exponential quadratic cost criterion. This type of estimators is well known as risk sensitive estimator in literature due to its origin from risk sensitive control law. The risk sensitive estimator is expected to have more robustness compare to its risk neutral counterpart. For linear Gaussian signal models, a closed-form solution exists and the risk-sensitive filter (RSF) has been proposed earlier [3]. However the limitations associated with EKF also inherited to ERSF. Similar to EKF, the ERSF also frequently diverges for highly nonlinear systems.

To overcome the drawbacks associated with ERSF several estimators have been proposed earlier by the present author. These are risk sensitive unscented Kalman filter (RSUKF) [4], central difference risk sensitive filter (CDRSF) [5], risk sensitive particle filter (RSPF) [6] and Adaptive grid risk sensitive filter (AGRSF) [7] etc. Among them in RSPF and AGRSF intractable integrals are evaluated using numerical techniques. So the computational load is very high for these two filters. On the other hand CDRSF and RSUKF are computationally efficient compared to RSPF and AGRSF.

Recently a new type of linear regression Kalman filter based on cubature and quadrature rule of integration has been proposed in literature [8]. The proposed filter is named as cubature Kalman filter (CKF) and expected to have better performance than EKF in spite of almost same computational burden. To overcome the limitations associated with ERSF a novel technique named as risk sensitive cubature Kalman filter (RSCKF) has been proposed in this paper. A simple, easy to follow derivation has been provided for approximate evaluation of intractable integrals using cubature and quadrature rules. To implement the proposed filter cubature points need to be generated propagated and updated. The final algorithm, as discussed later in this paper has similarities with UKF [9]. The proposed robust filter has been applied to a simple single dimensional problem which is severely nonlinear in nature. The simulation results for both single run as well as Monte Carlo run have been provided and compared with ERSF.

II. RISK SENSITIVE FILTERING

Consider the nonlinear plant described by the state and measurement equations as follows:

\[ x_{k+1} = \phi(x_k) + w_k \]  
\[ y_k = \gamma(x_k) + v_k \]

Where \( x_k \in \mathbb{R}^n \) denotes the state of the system, \( y_k \in \mathbb{R}^p \) is the measurement at the instance \( k \) where \( k = 0, 1, 2, 3, \ldots, N, \phi(x_k) \) and \( \gamma(x_k) \) are known nonlinear functions of \( x_k \) and \( k \). The process noise \( w_k \in \mathbb{R}^n \) and measurement noise \( v_k \in \mathbb{R}^p \) are assumed to be uncorrelated and normally distributed with \( Q_k \) and \( H_k \) respectively. The following notations have been used to represent the probability density functions

\[ f(x_{k+1}|x_k) \triangleq p_{x_{k+1}|x_k}(\cdot|x_k) \]
\[ g(y_k|x_k) \triangleq p_{y_k|x_k}(\cdot|x_k) \]

The objective is to estimate a known function \( \Phi(x) \) of the state variables. The estimate is designated as \( \hat{\Phi}(x) \) and its optimal value in the risk sensitive sense is denoted as \( \Phi^*(x) \) which
minimizes the cost function [1]

\[
C(\hat{\Phi}_1, \ldots, \hat{\Phi}_k) = E[\exp(\mu_1 \sum_{i=1}^{k-1} p_1(\Phi(x_i) - \hat{\Phi}_i)) + (\mu_2 p_2(\Phi(x_k) - \hat{\Phi}_k))]
\]

Where \( \mu_1 \geq 0 \) and \( \mu_2 > 0 \) are two risk parameters. Functions \( p_1(.) \) and \( p_2(.) \) are both strictly convex, continuous and bounded from below, attaining global minima at 0. In particular, the Minimum Risk Sensitive Estimate (MRSE) is defined by

\[
\hat{\Phi}_k^* = \arg \min \sigma_k(x_k)
\]

It can be shown that the MRSE satisfies the following recursion

\[
\sigma_k(x_k) = \int f(x_k|x_{k-1})g(y_k|x_k)\times \exp(\mu_1 p_1(\Phi(x_{k-1}) - \hat{\Phi}_{k-1}^*))\sigma_{k-1}(x_{k-1})dx_{k-1}
\]

\[
\hat{\Phi}_k^* = \arg \min_{\alpha \in R} \int \exp(\mu_2 p_2(\Phi(x_k) - \alpha))\sigma_k(x_k)dx_k
\]

Where \( \sigma_k(x_k) \) represents an information state [1] and may be normalized and \( \alpha \) is a parameter. If we assume the variables to be estimated are the state variables themselves (\( \Phi(x) = x \)), and both the convex functions \( p_1(.) \) and \( p_2(.) \) are known quadratic functions of vectors i.e., \( p_j(\epsilon) = \epsilon^T \epsilon \) for \( j = 1, 2 \); the equation (4) and (5) becomes

\[
\sigma_k(x_k) = \int f(x_k|x_{k-1})g(y_k|x_k)\exp[\mu_1(x_k - \hat{x}_{k|k-1}^T(x_k - \hat{x}_{k|k-1}))|\sigma_{k-1}(x_{k-1})dx_{k-1}
\]

\[
\hat{x}_{k|k}^* = \arg \min_{\alpha \in R} \int \exp(\mu_2 p_2(x_k - \alpha))\sigma_k(x_k)dx_k
\]

The above risk sensitive recursion is posterior in nature. Similar recursion for prior estimation is available in literature [6] and not mentioned here.

**III. FILTERING UNDER BAYESIAN FRAMEWORK**

In Bayesian estimation paradigm the states \( x_k \) is to be recursively calculated at time \( k \) taking different value of data \( y_{1:k} \) up to time \( k \). The prior probability density can be given by Chapman-Kolmogorov equation:

\[
p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|y_{1:k-1})dx_{k-1}
\]

The above equation is known as time update equation. The computation of posterior density function is done via Bayes rule.

\[
p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})}
\]

Where the normalizing constant

\[
p(y_k|y_{1:k-1}) = \int p(y_k|x_k)p(x_k|y_{1:k-1})dx_k
\]

Now for Gaussian noises and linear process and measurement the posterior and prior densities remain Gaussian in nature and the estimated value can be obtained optimally by celebrated Kalman filter. For Nonlinear system the probability density functions are no longer Gaussian in nature. But many times it is approximated as Gaussian and tried to find out the mean and covariance of prior as well as posterior density function.

**Time update:** The prior estimate is the mean of prior probability density function. So

\[
\hat{x}_{k|k-1} = E[|x_k|y_{1:k-1} = E[\phi(x_{k-1}) + w_k|y_{1:k-1}] = E[\phi(x_{k-1})|y_{1:k-1}]
\]

or,

\[
\hat{x}_{k|k-1} = \int \phi(x_{k-1})p(x_{k-1}|y_{1:k-1})dx_{k-1} = \int \phi(x_{k-1})N(x_{k-1}; \hat{x}_{k-1|k-1}, P_{k-1|k-1})dx_{k-1}
\]

\[
P_{k|k-1} = E[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T | y_{1:k-1}]
\]

\[
= \int \phi(x_{k-1})\phi^T(x_{k-1})N(x_{k-1}; \hat{x}_{k|k-1}, P_{k-1|k-1})dx_{k-1} - \hat{x}_{k|k-1}\hat{x}_{k|k-1}^T + Q_k
\]

**Measurement update:**

\[
p(y_k|y_{1:k-1}) = N(y_k; \hat{y}_{k|k-1}, P_{yy|k|k-1})
\]

where

\[
\hat{y}_{k|k-1} = \int \gamma(x_k)N(x_k; \hat{x}_{k|k-1}, P_{k|k-1})dx_k
\]

\[
P_{yy|k|k-1} = \int \gamma(x_k)\gamma^T(x_k)N(x_k; \hat{x}_{k|k-1}, P_{k|k-1})dx_k - \hat{y}_{k|k-1}\hat{y}_{k|k-1}^T
\]

\[
K_k = P_{yy|k|k-1}^{-1}P_{yy|k|k-1}^{-1}K_k
\]

Cross covariance

\[
P_{yy|k|k-1} = \int \gamma(x_k)\gamma(x_k)N(x_k; \hat{x}_{k|k-1}, P_{k|k-1})dx_k - \hat{x}_{k|k-1}\hat{y}_{k|k-1}^T
\]

On the receipt of new measurement \( y_k \) the posterior density

\[
p(x_k|y_{1:k}) = N(x_k; \hat{x}_{k|k}, P_k)
\]

where

\[
\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1})
\]

\[
P_{k|k} = P_{k|k-1} - K_kP_{yy|k|k-1}K_k^T
\]

From the equations above it is clear that to obtain state estimation, the above integrations need to be evaluated. Only for linear \( \phi(.) \) and \( \gamma(.) \) the closed form solution is available. For nonlinear functions the accuracy of the estimation depends on the accuracy of the approximate evaluation of the integrals.


IV. CUBATURE QUADRATURE EVALUATION OF MULTIDIMENSIONAL INTEGRAL

The intractable integrals described above can be evaluated using cubature quadrature rules of numerical integration.

**Theorem 1:** For an arbitrary function \( f(X), X \in \mathbb{R}^n \) the integral \( I(f) = \frac{1}{\sqrt{\mathcal{Z}^{(2n)n}}} \int_{\mathbb{R}^n} f(X) \exp(-\frac{1}{2}(X - \mu)^T \mathcal{Z}^{-1}(X - \mu)) dX \) can be expressed in spherical co-ordinate system as \( I(f) = \frac{1}{\sqrt{\mathcal{Z}^{(2n)n}}} \int_{U_n} |f(CrZ + \mu)| dz r^{n-1} e^{-r^2/2} dr \) where \( \mathcal{Z} = \mathcal{Z}^{(2n)n} \) is the Cholesky decomposition of \( \mathcal{Z} \), \( |Z| = 1 \), \( U_n \) is the surface of unit sphere.

**Proof:** Let us transform the integral \( I(f) \) to a spherical co-ordinate system \([10]\). Let \( X = CY + \mu \), \( \mathcal{Z} = CC^T \) is the Cholesky decomposition of \( \mathcal{Z} \). Then \( (X - \mu)^T \mathcal{Z}^{-1}(X - \mu) \) can be expressed in spherical co-ordinate system as \( |C| = 1 \), \( U_n \) is the surface of unit sphere.

Now let \( dZ \) at \( n \) dimensional space is zero mean unity variance \((13) \) can be approximately written as

\[
I(f) = \frac{1}{\sqrt{\mathcal{Z}^{(2n)n}}} \int_{R^n} f(X) e^{-\frac{1}{2}(X - \mu)^T \mathcal{Z}^{-1}(X - \mu)} dX
\]

or,

\[
I(f) = \frac{1}{\sqrt{\mathcal{Z}^{(2n)n}}} \int_{R^n} f(CY + \mu) e^{-\frac{1}{2}Y^T \mathcal{Z} Y} \sqrt{\mathcal{Z}} dY
\]

Now to integrate the rest of the term we would use the Gauss Laguerre quadrature formula. To cast the integration in that form we need some transformation. Let us assume \( t = \frac{1}{2}r^2 \).

With this transformation the equation \((15)\) becomes

\[
I(f) = \frac{1}{2^n \Gamma(n/2)} \int_{\mathbb{R}^n} f(\sqrt{2} \mu) e^{-t} dt
\]

Now the integration \( f(t) e^{-t} dt \) is approximated using first order Gauss Laguerre approximation with the node at \( n/2 \) and weight \( \Gamma(n/2) \) \([11]\) \([12], \text{pp. 34,131}\). So

\[
\int_{\mathbb{R}^n} f(t) e^{-t} dt \approx \Gamma(n/2) f(n/2)
\]

Substituting \((17)\) in \((16)\) we get

\[
I(f) = \frac{1}{2^n \Gamma(n/2)} \sum_{i=1}^{2n} f(\sqrt{2} \mu_i) e^{-\frac{r^2}{2}}
\]

The cubature rule used above is third degree fully symmetric cubature rule. If we consider zero mean unity variance \((13)\) can be approximately written as \([8]\)

\[
\int_{U_n} f(rZ) dz \approx \frac{2\sqrt{\pi}}{2n \Gamma(n/2)} \sum_{i=1}^{2n} f(\sqrt{n} \mu_i)
\]

V. FORMULATION OF RSCKF

To obtain the risk sensitive posterior estimation, closed form evaluation of the integral described in \((6)\) is necessary. But the closed form solution only exists for linear Gaussian system.
For nonlinear system the integral is intractable. Here cubature based numerical integration method is used to calculate the integral approximately. As it can be easily understood that when the equation (6) is approximated using cubature rule the risk sensitive estimator obtained will be posterior type. Similar type of recursive formulae are available for prior risk sensitive estimation when approximated using cubature rule prior RSCKF would be obtained.

The posterior recursive form of risk sensitive estimation can be described using equation (6) and (7). Let us introduce a new distribution

\[\sigma_{k-1|k-1}(x_k) = \exp(\mu_1(x_k - \hat{x}_{k|k-1})^T \times (x_k - \hat{x}_{k|k-1}) \sigma_{k-1|k-1}(x_k-1) \]  

(19)

Hence

\[\sigma_{k|k-1}(x_k-1) = \int f(x_k|x_k-1) \sigma_{k|k-1}(x_k-1) dx_{k-1} \]  

(20)

\[\sigma_{k|k}(x_k) = g(y_k|x_k) \sigma_{k|k-1}(x_k-1) \]  

(21)

Equation (19) may be considered as risk sensitive update step. Equation (20) and (21) may be considered as time update and measurement update respectively. It should be noted here that, as in cubature Kalman filter Gaussian approximation is maintained, the optimal estimate obtained from (7) is simply the mean value of the distribution.

Summarizing all the points mentioned above, the computing steps can be categorized as follows:

A. General RSCKF algorithm for nonlinear systems

Step (i) Filter initialization
- Initialize the filter with \(x_{0|0} \) and \( \hat{P}_{0|0} \)
- Generate cubature points \( \xi_i = \sqrt{n}[i] \) where \( i = 1, 2, ..., 2n \) \( n \) is the dimension of state vector.
- Generate weights \( w = 1/n \)

Step (ii) Predictor step
- Perform Cholesky factorization of posterior error covariance

\[P_{k|k} = S_{k|k} S_{k|k}^T \]

- Evaluate cubature points \( \chi_{i,k|k} = \hat{x}_{k|k} + \sigma_{k|k} \xi_i \), where \( i = 1, 2, ..., 2n \)
- Update cubature points

\[\chi_{i,k+1|k} = \phi(\chi_{i,k|k}) \]

- Compute time updated mean and covariance

\[\sigma_{k+1|k} = \sum_{i=1}^{2n} \chi_{i,k+1|k} \]

\[P_{k+1|k} = \sum_{i=1}^{2n} \chi_{i,k+1|k} - \sigma_{k+1|k} \sum_{i=1}^{2n} \chi_{i,k+1|k} - \sigma_{k+1|k} + Q_k \]

- The risk sensitive mean remains the same and the risk sensitive covariance is updated with

\[P_{k+1|k}^+ = (P_{k+1|k} - 2\mu_1 I)^{-1} \]

Step (iii) Corrector step or measurement update
- Perform Cholesky factorization of prior error covariance

\[P_{k+1|k} = S_{k+1|k} S_{k+1|k}^T \]

- Evaluate cubature points \( \chi_{i,k+1|k} = S_{k+1|k} \xi_i + \hat{x}_{k+1|k} \), where \( i = 1, 2, ..., 2n \)
- Projected measurements at each cubature points

\[Y_{i,k+1|k} = \gamma(\chi_{i,k+1|k}) \]

- Estimate the predicted measurement

\[\hat{y}_{k+1} = \sum_{i=1}^{2n} Y_{i,k+1|k} \]

- Calculation of covariances

\[P_{y+k+1|y+k+1} = \sum_{i=1}^{2n} [Y_{i,k+1|k} - \hat{y}_{k+1}][Y_{i,k+1|k} - \hat{y}_{k+1}]^T + R_k \]

\[P_{x+k+1|y+k+1} = \sum_{i=1}^{2n} [\chi_{i,k+1|k} - \sigma_{k+1|k}][Y_{i,k+1|k} - \hat{y}_{k+1}]^T \]

- Calculation of Kalman gain

\[K_{k+1} = P_{x+k+1|y+k+1} P_{y+k+1|y+k+1}^{-1} \]

- Posterior state values

\[\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}(y_{k+1} - \hat{y}_{k+1}) \]

- Posterior error covariance matrix is given by

\[P_{k+1|k+1} = P_{k+1|k} - K_{k+1} P_{y+k+1|y+k+1} K_{k+1}^T \]

Step (iv) Optimal risk sensitive estimate
- With Gaussian assumption the optimal posterior risk sensitive estimate would be

\[\hat{x}_{k+1|k+1} = \sigma_{k+1|k+1} \]

Remarks:
1) Compare to extended risk sensitive filter (ERSF), the proposed filter is derivative free i.e. to implement it neither Jacobian nor Hessian matrix need to be calculated. This may be considered as an added advantage from the computation point of view.
2) To implement the proposed RSCKF at each step \(2n(\text{n is the dimension of state vector})\) number of cubature points need to be calculated. So the computational burden increases linearly with the dimensions. Also the cubature points and weights can be calculated off-line.
3) The above described algorithm is to calculate posterior risk sensitive estimation of states. Similar algorithm can be developed for prior estimation of states using RSCKF.
4) The condition \((P_{k+1|k} - 2\mu_1 I)^{-1} > 0\) need to be satisfied at each step for risk sensitive error covariance to be positive definite. This condition limits the upper value of \(\mu_1\).
leading to a track loss situation. The estimate to settle down at the wrong equilibrium point, becomes challenging because moderate estimation error forces either of the two stable equilibrium points. The problem stable. In the absence of any bias, the system hovers around one at the origin is unstable and the other two are at and

d
in figure 2. The figure reveals that RMSE of ERSF diverges 1000 Monte Carlo run for both ERSF and RSCKF and shown in comparison to the traditional ERSF.

Improvement of estimation accuracy with the proposed filter whereas for RSCKF it is about 7%. The numbers indicate the percentage fail count is about 26% estimator settles on +1 equilibrium point out of 100 Monte Carlo run. For ERSF the percentage fail count is about 26% whereas for RSCKF it is about 7%. The numbers indicate the improvement of estimation accuracy with the proposed filter in comparison to the traditional ERSF.

Root mean square error (RMSE) has been carried out for 1000 Monte Carlo run for both ERSF and RSCKF and shown in figure 2. The figure reveals that RMSE of ERSF diverges (due to large population of track loss cases), whereas the RMSE of RSCKF settles approximately at 0.5. It should also be noted that the computational cost of both the filters are almost same.

VII. DISCUSSIONS AND CONCLUSION

Solution of nonlinear risk sensitive estimation problem has been proposed using cubature quadrature rule of integration. The theory and algorithm of proposed filter named as risk sensitive cubature Kalman filter (RSCKF) has been developed. Also a derivation for approximate evaluation of intractable integrals using cubature and quadrature rules has been provided. With the help of an example it has been shown that the proposed filter can overcome the draw back associated with ERSF. The algorithm of RSCKF has come out similar to RSUKF. The detail comparison of proposed filter with RSUKF and square root version of it remain under the scope of future work. Due to numerical efficiency and expected robustness, the proposed filter may find place in on-board real life applications.

REFERENCES


