Nonlinear Estimation Using Transformed Gauss-Hermite Quadrature Points

Abhinoy Kumar Singh
Department of Electrical Engineering
Indian Institute of Technology Patna
Patna, Bihar, India
Email:abhinoy@iitp.ac.in

Shovan Bhaumik
Department of Electrical Engineering
Indian Institute of Technology Patna
Patna, Bihar, India
Email:shovan.bhaumik@iitp.ac.in
Telephone: +91 612 255 2049
Fax: +91-612-227 7383

***Abstract*—An ongoing work proposing a new method for nonlinear filtering problem is reported. The intractable integrals have appeared in nonlinear estimation problem been approximately evaluated using Gauss-Hermite quadrature rule. An orthogonal transformation has been applied on Gauss-Hermite points in order to obtain more accurate estimation for higher order problems. The developed method is named transformed Gauss-Hermite filter. The efficacy of proposed filter compared to ordinary Gauss-Hermite filter is demonstrated with the help of an example.**

***Index Terms*—Nonlinear filtering, Gauss-Hermite quadrature rule, Orthogonal transformation.**

I. INTRODUCTION

There are many practical and industrial problems where we need to estimate states for nonlinear systems. With the increasing demands from the practitioners to develop more accurate and computationally efficient algorithm for nonlinear non Gaussian system, various filtering methods have been proposed.

In Bayesian framework, to obtain state estimation, one has to solve an integral which is intractable for nonlinear systems. Initially the extended Kalman filter (EKF) [1][2] is introduced where the process and measurement equations are linearized on their respective current mean value. The estimated mean and covariance are propagated and updated recursively. EKF failed to attract the designers due to its well reported limitations [1][2] including smoothness requirement of functions, noise Gaussian restriction and lack of convergence for highly nonlinear systems. As a result several numerical filtering techniques namely the unscented Kalman filter (UKF) [3], the Gauss-Hermit filter (GHF) [4][5], the central difference filter (CDF) [6], *etc.* have been evolved in literature. In the above mentioned algorithms the intractable integrals arise in nonlinear or non Gaussian problems have been approximately solved using deterministic sample points, and weights associated with them.

In the GHF, the intractable integrals have been approximately evaluated using Gauss-Hermite quadrature. Although quadrature rule has been developed for solving 1D integrals, it could easily be extended for multidimensional cases using product rule. Accuracy of GHF depends on the number of quadrature points used to evaluate the integrals. The number of quadrature points required to solve the integrals increases exponentially with the dimension of system, hence the GHF suffers from *the curse of dimensionality* problem.

In this paper, we generate a new set of sample points by performing an orthogonal transformation on multidimensional Gauss-Hermite quadrature points. The weights associated with the new sample points remain same. The transformed points are used to approximate the posterior probability density function as Gaussian. The proposed nonlinear estimator is named transformed Gauss-Hermite filter (TGHF). It shows improvement in performance compared to ordinary GHF for higher dimensional systems.

The paper is organized as follows: Section 2 explains about the Baysian framework for filtering. This is followed by Gauss-Hermite filtering basics. Section 4 describes orthogonal transformation on Gauss-Hermite points. Simulation results are presented in section 5 and concluding remarks are placed in section 6.

II. FILTERING UNDER BAYESIAN FRAMEWORK

Let us consider a nonlinear plant described by the process and measurement equations as follows:

$$x_k = \phi(x_{k-1}) + \eta_{k-1} \tag{1}$$

$$y_k = \gamma(x_k) + v_k \tag{2}$$

where $x_k \in \mathbb{R}^n$ denotes the state of the system, $y_k \in \mathbb{R}^p$ is the measurement at the instant k where $k = \{0, 1, 2, 3, ...N\}$, $\phi(x_k)$ and $\gamma(x_k)$ are known nonlinear functions of x_k and k. The process noise $\eta_k \in \mathbb{R}^n$ and measurement noise $v_k \in \mathbb{R}^p$ are assumed to be uncorrelated, normally distributed with zero mean and covariance Q_k and R_k respectively.

In Bayesian estimation paradigm, the states x_k is to be estimated recursively at time k considering measurement data, $y_{1:k}$, up to time k. The prior probability density can be given by Chapman-Kolmogorov equation

$$p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|y_{1:k-1})dx_{k-1} \tag{3}$$

978-1-4673-6190-3/13/$31.00 © 2013 IEEE

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

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})} \tag{4}$$

where the normalizing constant

$$p(y_k|y_{1:k-1}) = \int p(y_k|x_k)p(x_k|y_{1:k-1})dx_k \tag{5}$$

For linear Gaussian system, the posterior and prior densities remain Gaussian in nature and the estimated values can be obtained optimally by the celebrated Kalman filter. Although for nonlinear system, the density functions are no longer Gaussian in nature, many times it is approximated as Gaussian and subsequently mean and covariance of prior as well as posterior density functions are evaluated.

From the above discussions, it is clear that to obtain estimated states, the integrals (3) and (4) need to be evaluated. As the integrals can not be solved analytically, several numerical integration techniques [7] have been proposed. Further, the accuracy of the estimation depends on the accuracy of the approximate evaluation of the integrals. In GHF formulation, intractable integral is approximately evaluated using Gauss-Hermite quadrature points and weights associated with them.

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

$$\begin{aligned}\hat{x}_{k|k-1} &= E[x_k|y_{1:k-1}] \\ &= E[(\phi(x_{k-1}) + \eta_{k-1})|y_{1:k-1}] \\ &= E[\phi(x_{k-1})|y_{1:k-1}]\end{aligned}$$

or,

$$\begin{aligned}\hat{x}_{k|k-1} &= \int \phi(x_{k-1})p(x_{k-1}|y_{1:k-1})dx_{k-1} \\ &= \int \phi(x_{k-1})\aleph(x_{k-1};\hat{x}_{k-1|k-1},P_{k-1|k-1})dx_{k-1}\end{aligned}$$

$$\begin{aligned}P_{k|k-1} &= E[(x_k-\hat{x}_{k|k-1})(x_k-\hat{x}_{k|k-1})^T|y_{1:k-1}] \\ &= \int [\phi(x_{k-1})\phi^T(x_{k-1})\aleph(x_{k-1};\hat{x}_{k-1|k-1},P_{k-1|k-1}) \\ &\quad dx_{k-1}] - \hat{x}_{k|k-1}\hat{x}_{k|k-1}^T + Q_k\end{aligned}$$

Measurement update: The measurement update under the approximation of Gaussian posteriori pdf is given by,

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

where

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

$$\begin{aligned}P_{yy,k|k-1} &= \int \gamma(x_k)\gamma^T(x_k)\aleph(x_k;\hat{x}_{k|k-1},P_{k|k-1})dx_k \\ &\quad - \hat{y}_{k|k-1}\hat{y}_{k|k-1}^T + R_k\end{aligned}$$

Cross covariance

$$\begin{aligned}P_{xy,k|k-1} &= \int x_k\gamma^T(x_k)\aleph(x_k;\hat{x}_{k|k-1},P_{k|k-1})dx_k \\ &\quad - \hat{x}_{k|k-1}\hat{y}_{k|k-1}^T\end{aligned}$$

On the receipt of new measurement y_k, the posterior density

$$p(x_k|y_{1:k}) = \aleph(x_k;\hat{x}_{k|k},P_{k|k})$$

where

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1}) \\ P_{k|k} &= P_{k|k-1} - K_kP_{yy,k|k-1}K_k^T \\ K_k &= P_{xy,k|k-1}P_{yy,k|k-1}^{-1}\end{aligned}$$

III. Gauss-Hermite Filter

The basic principle involved in GHF is Gauss-Hermite quadrature rule of integration which provides an approximate way to solve the intractable integrals encountered in nonlinear Bayesian filtering framework. Although the Gauss-Hermite quadrature rule of integration is available in literature [8][9] for more than fifty years, the same has been incorporated in estimation very recently, mainly due to work of Ito and Xiong [6]. In Gauss-Hermite filter, unknown probability density function (pdf) has been approximated as Gaussian using a set of Gauss-Hermite quadrature points and their respective weights.

An integral of any function $f(x)$ in the form of

$$I = \int_{-\alpha}^{\alpha} f(x)e^{-x^2}dx \tag{6}$$

may be numerically evaluated using N quadrature points

$$I \approx \sum_{i=1}^{N} f(q_i)w_i \tag{7}$$

Where q_i and w_i are the i_{th} quadrature point and it's corresponding weight.

To calculate the quadrature points and the weights, let us consider a symmetric tridiagonal matrix J having zero diagonal elements and $J_{i,i+1} = \sqrt{i/2}; 1 \le i \le (N-1)$. The quadrature points are at $q_i = \sqrt{2}\psi_i$, where ψ_i is the i_{th} eigenvalue of the matrix J. The i_{th} weight, w_i, is chosen as $w_i = \kappa_{i1}^2$, where κ_{i1} is the first element of the i_{th} normalized eigenvector of J [4][5].

The single dimensional quadrature formula could be extended for multidimensional problem by making use of the product rule. Using product rule, n dimensional integral,

$$I_N = \int_{R_n} f(s)\frac{1}{(2\pi)^{n/2}}e^{-(1/2)|s|^2}ds$$

could be approximately evaluated as

$$I_N \approx \sum_{i_1}^{N}...\sum_{i_n}^{N} f(q_{i_1},q_{i_2},...,q_{i_n})w_{i_1}w_{i_2}...w_{i_n}$$

Evaluation of expected value for N-point GHF in n dimensional system requires N^n number of quadrature points and corresponding weights. For an example, for a three dimensional system and three point GHF, twenty-seven quadrature points and weights are required which may be expressed as $\{q_i,q_j,q_k\}$ and $\{w_iw_jw_k\}$ respectively for $i = 1,2,3;$

$j = 1, 2, 3$; and $k = 1, 2, 3$. As the number of quadrature points increases exponentially with increase in dimension of the systems, the GHF suffers from *the curse of dimensionality* problem. To a very limited extent this could be overcome by ignoring the quadrature points on the diagonal because weights associated with them are very small, hence they contribute negligible to the computation of the integrals. In another approach, Bin Jia [10] proposed to use Smolyak rule instead of product rule for multidimensional Gauss-Hermite integration. This considerably decreases quadrature points requirement for multidimensional integrals. In this work, we adopt the product rule and all the quadrature points are retained.

IV. Transformed Gauss-Hermite filter

In the previous section, the methodology for generating quadrature points and their corresponding weights has been discussed. In this section we present an orthogonal transformation matrix [11] with which we transform all the Gauss-Hermite points to obtain a new set of quadrature points. It is observed that more accurate estimation (particularly in higher dimensional systems) is obtained with transformed points compared to simple Gauss-Hermite points. The reason behind the improvement in accuracy with described orthogonal transformation could be smaller higher order terms of Taylor series for high dimensional system.

After we apply orthogonal transformation on the quadrature points, the points re-orient on n dimensional space. The $n \times n$ orthogonal transformation matrix ρ is taken as $\rho = [\rho_1, \rho_2, ..., \rho_n]$, where $\rho_m = [\rho_{m,1}, \rho_{m,2}, ..., \rho_{m,n}]^T$ and the element of ρ_m is given by [11] $\rho_{m,2r-1} = \sqrt{\frac{2}{n}} cos((2r-1)m\pi/n)$, $\rho_{m,2r} = \sqrt{\frac{2}{n}} sin((2r-1)m\pi/n)$, $r = 1, 2, ..., [n/2]$. if n is odd then $\rho_{m,n} = (-1)^k/\sqrt{n}$, where $[n/2]$ is the greatest integer not exceeding $n/2$.

New sample points after transformation are

$$\xi = \rho \times q \tag{8}$$

The dimension of the orthogonal matrix ρ is $n \times n$ and the quadrature points q is $n \times N^n$. Hence dimension of the new sample points will be $n \times N^n$. The algorithm of the proposed TGHF is same as the GHF, except the Gauss-Hermite quadrature points are replaced by transformed quadrature points.

V. Simulation Result

The proposed estimator has been applied to estimate states of the following nonlinear systems

Process equation:

$$x_k = 20cos(x_{k-1}) + \eta_{k-1} \tag{9}$$

Measurement equation:

$$y_k = \sqrt{1 + x_k^T x_k} + v_k \tag{10}$$

Here x_k is an n dimensional Gaussian random variable which represent the value of states at k_{th} step. y_k is the measurement at k_{th} step. η_k is the Gaussian process noise. $\eta_k \sim \aleph(0_{n\times1}, I_n)$, where $0_{n\times1}$ denotes an n dimensional column matrix having all the elements equal to zero and I_n is an identity matrix of order n. $v_k \sim \aleph(0, 1)$ is Gaussian measurement noise.

The initial truth states are considered as $x_0 = 0.1 \times 0_{n\times1}$. The filter is initialized with a value of $\hat{x}_0$ and P_0, where $\hat{x}_0 = 0_{n\times1}$ and $P_0 = I_n$. The states are estimated using the GHF and the proposed TGHF filter. Root mean square error (RMSE) has been calculated with the GHF and the TGHF. For M number of Monte Carlo runs RMSE at k_{th} step is defined as:

$$RMSE_k = \sqrt{\frac{1}{M}\sum_{i=1}^{M} e_{i,k}^2} \tag{11}$$

where the error $e_{i,k}$ is given as: $e_{i,k} = (x_{i,k} - \hat{x}_{i,k})$, $i = 1, 2, ..., M$.

For a ten dimensional system ($n = 10$) described by equations (9) and (10), the RMSE of first state variable obtained from the GHF and TGHF using 50 Monte Carlo runs have been plotted in Fig 1. Both the filters use 3^{10} number of quadrature points in ten dimensional space to calculate the mean and covariance. Fig 1 reveals that RMSE obtained from the proposed TGHF is significantly low compared to ordinary GHF. Similar results are obtained for other state variables hence they are omitted here.

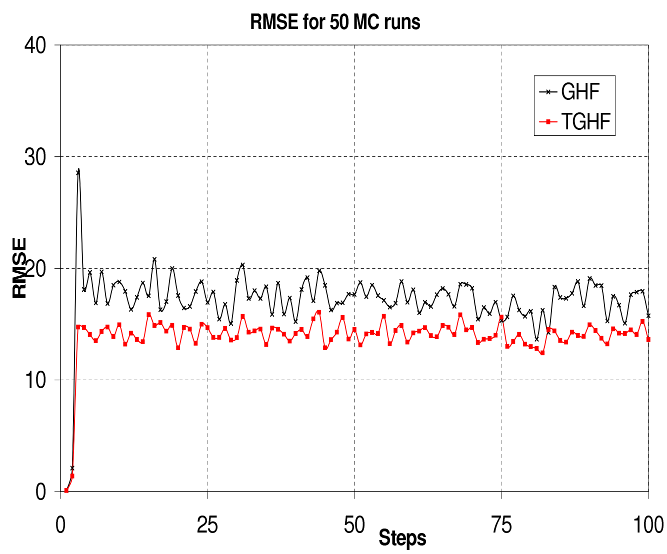

Fig. 1. RMSE plot of 50 Monte Carlo run for $n = 10$.

We also calculate the averaged RMSE value over 100 time steps using the GHF and the TGHF for different n ranging from 2 to 10. The RMSE values averaged over time are tabulated in Table 1. From the Table 1 it is clear that the performance of the TGHF is better compared to ordinary GHF. It can also be seen from Table 1 that as dimension of the system goes high, the TGHF performs much better than traditional GHF.

TABLE I
AVERAGED RMSE

Dimension (n)	**GHF**	**TGHF**
2	17.4565	17.5112
3	17.0160	14.8745
5	17.1044	14.2124
10	17.1520	13.9221

VI. DISCUSSIONS AND CONCLUSION

In this paper, a nonlinear filtering method based on orthogonal transformation of Gauss-Hermite quadrature points has been proposed. The proposed method is named transformed Gauss-Hermite filter. In an example, compared to ordinary GHF, the proposed filter demonstrates higher accuracy for estimating states of higher order systems. However the major limitation of TGHF which has been inherited from GHF is exponential increase in computation cost with dimension. Research work is going to overcome *the curse of dimensionality* problem encountered in TGHF. Due to enhanced accuracy, the proposed filter may replace ordinary GHF for nonlinear estimation problem.

REFERENCES

[1] Brian D.O. Anderson, and John B. Moore, “Optimal filtering”, Dover Publication, INC., Mineola, New York, 2005.

[2] Y. Bar-Shalom, X. Rong Li, and T. Kirubarajan, “Estimation with application to tracking and navigation”, John Wiley and Sons, INC., New York, 2001.

[3] Simon Julier, Jeffrey Uhlmann, and Hugh F. Durrant-Whyte, “A new method for the nonlinear transformation of means and covariances in filters and estimators”, IEEE Transactions on Automatic Control, vol. 45, no. 3, pp. 477-482, March 2000.

[4] Arasaratnam I., Simon Haykin, and Robert J. Elliott, “Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature”, Proceedings of the IEEE, vol. 95, no. 5, pp. 953-977, May 2007.

[5] Gautam Chalasani, and Shovan Bhaumik, “Bearing only tracking using Gauss-Hermite filter”, Proceedings, IEEE conference on Industrial Electronics and Applications (ICIEA), pp. 1549-1554, 2012.

[6] K. Ito, and K. Xiong, “Gaussian filters for nonlinear filtering problems”, IEEE Transaction on Automatic Control, vol. 45, no. 5, pp. 910-927, May 2000.

[7] Y. Wu, D. Hu, M. Wu, and X. Hu, “A numerical-integration perspective on Gaussian filters” IEEE Transactions on Signal Processing, vol. 54, no. 8, pp. 2910-2921, 2006.

[8] F.B. Hildebrand, “Approximate calculation of integrals”, 2nd ed., New York, Dover Publication, 2008.

[9] V.I. Krylov, “Approximate calculation of integrals”, NC Mineola, New York, Dover Publication, 2005.

[10] Bin Jia , Ming Xin, and Yang Cheng, “Sparse-grid quadrature nonlinear filtering”, Automatica, vol. 48, no. 2, pp. 327-341, Feb. 2012.

[11] Lubin Chang, Baiqing Hu, An Li, and Fangjun Qin, “Transformed unscented Kalman filter”, IEEE Transaction on Automatic Control, vol. 58, no. 1, pp. 252-257, Jan. 2013.