Performance Analysis of Derivative-Free Filters

Jindřich Duník, Miroslav Šimandl, Ondřej Straka, Ladislav Král

Research Centre Data - Algorithms - Decision Making
Department of Cybernetics
Faculty of Applied Sciences
University of West Bohemia in Pilsen
Czech Republic

December 12, 2005

Description of System

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k, \ \ k = 0, 1, 2, \ldots$$
$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \ \ k = 0, 1, 2, \ldots$$

Filtering Problem

The aim of the filtering is to find the probability density function (pdf) of the state $\mathbf{x}_k$ conditioned by the measurements $\mathbf{z}^k = [\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_k]$.

$$p(\mathbf{x}_k | \mathbf{z}^k) = ?$$

Solution of the Filtering Problem

Bayesian Recursive Relations (BRR's)

Solution of the filtering problem is given by the BRR's

$$p(\mathbf{x}_k|\mathbf{z}^k) = \frac{p(\mathbf{x}_k|\mathbf{z}^{k-1})p(\mathbf{z}_k|\mathbf{x}_k)}{\int p(\mathbf{x}_k|\mathbf{z}^{k-1})p(\mathbf{z}_k|\mathbf{x}_k)d\mathbf{x}_k},$$

$$p(\mathbf{x}_{k+1}|\mathbf{z}^k) = \int p(\mathbf{x}_k|\mathbf{z}^k)p(\mathbf{x}_{k+1}|\mathbf{x}_k)d\mathbf{x}_k,$$

where $p(\mathbf{x}_0|\mathbf{z}^{-1}) = p(\mathbf{x}_0)$.

Solution of the BRR's

- exact solution
- approximative solution
  - local methods
  - global methods

Key Aspect of Local Methods

The local methods are based on a suitable approximation of the system description so that the technique of the Kalman Filter design can be used in the area of nonlinear systems.

Resultant estimates: described by a mean and a covariance matrix.

Advant. and Disadvant.

Advantage is

- simplicity of the solution of the BRR's.

Disadvantage is

- impossibility to ensure the convergence of the state estimate.

Approaches in Local Estimation

- Standard approach, e.g.

  Extended Kalman Filter, Second Order Filter

- Novel derivative-free approach, e.g.

  Unscented Kalman Filter,

  (Julier, et al., 2000)

  Divide Difference Filter

  (Nørgaard, et al., 2000)

Key Aspect of Global Methods

The global methods are mainly based on an appropriate approximation of the description of the pdf's.
Resultant estimates: in form of a general pdf.

Advantages and Disadvantages

Advantage is

- the substantial improvement of the estimation quality in comparison with the local methods.

Disadvantage is

- growth of the computational demands in comparison with the local methods.

Approaches in Global Estimation

- Simulation approach, e.g. Particle Filter,
  - (Liu and Chen, 1998) - sequential Monte Carlo methods for dynamic systems,
  - (Doucet, et al., 2001) - Monte Carlo methods in practise,
  - (van der Merwe and Wan, 2003) - design of the Unscented Particle Filter based on the Unscented Kalman Filters.
- Numerical approach, e.g. Point-Mass Filter,
  - (Bucy and Senne, 1971) - digital synthesis of nonlinear filters,
  - (Kramer and Sorenson, 1988) - piece-wise constant approximation,
  - (Šimandl, et al., 2002) - anticipative design in point-mass approach.
- Analytical approach, e.g. Gaussian Sum Filter,
  - (Sorenson, 1974) - on development of practical filters,
  - (Anderson, 1985) - adaptive forgetting in recursive identification,
  - (Šimandl and Duník, 2005) - utilization of the novel derivative-free local filters in the Gaussian sum framework to design the Sigma Point Gaussian Sum Filter.

System Specification

Nonlinear Non-Gaussian System (van der Merwe, Wan (2003))

$$x_{k+1} = 0.5x_k + 1 + \sin(0.04\pi k) + w_k, k = 0, 1, \ldots, 60,$$

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

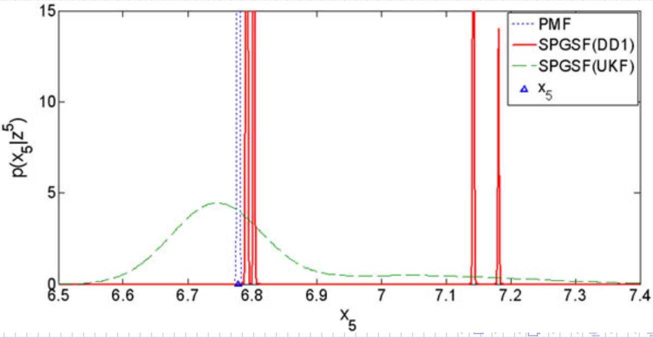
where

- $p(x_0|z^{-1}) = p(x_0) = \sum_{j=1}^{5} 0.2 \times \mathcal{N}(x_0 : j-3, 10)$,
- $p(w_k) = Ga(3,2) \approx \hat{p}(w_k) = 0.29 \times \mathcal{N}(w_k : 2.14, 0.72) + {} + 0.18 \times \mathcal{N}(w_k : 7.45, 8.05) + {} + 0.53 \times \mathcal{N}(w_k : 4.31, 2.29), \forall k$,
- $p(v_k) = \mathcal{N}(v_k : 0, 10^{-5}), \forall k$.

Experimental Results

Quality of Filter's Point Estimates

	SPGSF(DD1)	SPGSF(UKF)	PMF
MSE	0.0222	0.0133	$3.55 \cdot 10^{-4}$
Time (s)	2.520	2.140	3170



Issue of Novel Global Filtering Methods

- Recently several novel global methods based on the novel local derivative-free methods have been proposed, e.g. the Sigma Point Gaussian Sum Filter (SPGSF) (Šimandl and Duník, 2005) or the Unscented Particle Filter (van der Merwe and Wan, 2003), .
- Although, the comparison of these novel global methods should be firstly performed from the viewpoint of the estimated pdf's, they have been compared from the viewpoint of the point estimates only.
- Moreover, the estimated pdf's by the particular SPGSF's could have totally different and unexpected shapes.

Goal of Paper

- To find out the reason of the quite different shapes of the estimated pdf's of the particular Sigma Point Gaussian Sum Filters.

Issue of Novel Global Filtering Methods

- Recently several novel global methods based on the novel local derivative-free methods have been proposed, e.g. the Sigma Point Gaussian Sum Filter (SPGSF) (Šimandl and Duník, 2005) or the Unscented Particle Filter (van der Merwe and Wan, 2003), .
- Although, the comparison of these novel global methods should be firstly performed from the viewpoint of the estimated pdf's, they have been compared from the viewpoint of the point estimates only.
- Moreover, the estimated pdf's by the particular SPGSF's could have totally different and unexpected shapes.

Goal of Paper

- To find out the reason of the quite different shapes of the estimated pdf's of the particular Sigma Point Gaussian Sum Filters.

Review of (Sigma Point) Gaussian Sum Algorithm

- **Initialization:** Initial condition is assumed in the form

$$p(\mathbf{x}_0|\mathbf{z}^{-1}) = \sum_{j=1}^{N_0'} \alpha_{j,k} \mathcal{N}\{\mathbf{x}_{j,0} : \hat{\mathbf{x}}_{j,0}^{'}, \mathbf{P}_{j,0}^{'}\}.$$

- **Filtering step:** Multiple application of the filtering part of a local filter for each pair $\hat{\mathbf{x}}_{j,k}^{'}$ and $\mathbf{P}_{j,k}^{'}$ leads to the filtering pdf

$$p(\mathbf{x}_k|\mathbf{z}^k) \approx \sum_{j=1}^{N_k} \beta_{j,k} \mathcal{N}\{\mathbf{x}_{j,k} : \hat{\mathbf{x}}_{j,k}, \mathbf{P}_{j,k}\}.$$

- **Reduction of terms in the mixture of Gaussians.**
- **Prediction step:** Multiple application of the prediction part of a local filter for each pair $\hat{\mathbf{x}}_{j,k}$ and $\mathbf{P}_{j,k}$ leads to the prediction pdf

$$p(\mathbf{x}_{k+1}|\mathbf{z}^k) \approx \sum_{j=1}^{N_k'} \gamma_{j,k} \mathcal{N}\{\mathbf{x}_{j,k+1} : \hat{\mathbf{x}}_{j,k+1}^{'}, \mathbf{P}_{j,k+1}^{'}\}.$$

Transformation of Random Variable in Local Filters

- The crucial feature of a local filter is the manner of transformation of a random variable via a nonlinear function.
- Consider the random variables $\mathbf{x}$, $\mathbf{y}$ which are related through the nonlinear function $\mathbf{y} = \mathbf{g}(\mathbf{x})$.
- The random variable $\mathbf{x}$ is given by the first two moments, i.e. by
  - the mean $\bar{\mathbf{x}}$
  - and the covariance matrix $\mathbf{P}_x$.
- The aim is to compute the characteristics of the random variable $\mathbf{y}$, i.e.
  - the mean $\bar{\mathbf{y}} = E[\mathbf{y}]$,
  - the covariance matrix $\mathbf{P}_y = E[(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^T]$,
  - and the cross-covariance matrix $\mathbf{P}_{xy} = E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{y} - \bar{\mathbf{y}})^T]$.

Divide Difference Filter 1st Order (DD1) (Nørgaard, et al., 2000)

- The DD1 is based on the approximation of the nonlinear function $g(\cdot)$ by means of the Stirling's Interpolation 1st Order, i.e.

$$y = g(x) \approx g^{DD1}(x) = g(\bar{x}) + \frac{g(\bar{x} + h\sqrt{P_x}) - g(\bar{x} - h\sqrt{P_x})}{2h}\Delta x,$$

  where $\Delta x = x - \bar{x}$.

- The approximative characteristics of the random variable y can be calculated in accord with
  - $\bar{y}_A^{DD1} = g(\bar{x})$,
  - $P_{y,A}^{DD1} = \frac{1}{4h^2}\left(g(\bar{x} + h\sqrt{P_x}) - g(\bar{x} - h\sqrt{P_x})\right)^2$,
  - $P_{xy,A}^{DD1} = \frac{\sqrt{P_x}}{2h}\left(g(\bar{x} + h\sqrt{P_x}) - g(\bar{x} - h\sqrt{P_x})\right)$.

Divide Difference Filter 2nd Order (DD2) (Nørgaard, et al., 2000)

- The DD2 utilizes the more exact approximation of the nonlinear function $g(\cdot)$ based on the Stirling's Interpolation 2nd Order which is

$$y = g(x) \approx g^{DD1}(x) + \frac{g(\bar{x} + h\sqrt{P_x}) + g(\bar{x} - h\sqrt{P_x}) - 2g(\bar{x})}{2h^2} (\Delta x)^2 ,$$

where $\Delta x = x - \bar{x}$.

- The approximative characteristics of the random variable y can be calculated in accord with
  - $\bar{y}_A^{DD2} = \frac{h^2-1}{h^2} g(\bar{x}) + \frac{1}{2h^2} \left(g(\bar{x} + h\sqrt{P_x}) + g(\bar{x} - h\sqrt{P_x}) \right),$
  - $P_{y,A}^{DD2} = P_{y,A}^{DD1} + \frac{h^2-1}{4h^4} \left(g(\bar{x} + h\sqrt{P_x}) + g(\bar{x} - h\sqrt{P_x}) - 2g(\bar{x}) \right)^2 = = P_{y,A}^{DD1} + P_{y,e}^{DD2},$
  - $P_{xy,A}^{DD2} = \frac{\sqrt{P_x}}{2h} \left(g(\bar{x} + h\sqrt{P_x}) - g(\bar{x} - h\sqrt{P_x}) \right) = P_{xy,A}^{DD1}.$
- The accrual $P_{y,e}^{DD2}$ becomes significant especially for "highly" nonlinear function $g(\cdot)$.

Unscented Kalman Filter (UKF) (Julier, et al., 2000)

- The UKF is based on the Unscented Transformation, where the random variable x is approximated by the set of deterministically chosen weighted points (σ-points). The set is computed as
  - $\mathcal{X}_0 = \bar{\mathbf{x}}, \mathcal{W}_0 = \frac{\kappa}{1+\kappa}$,
  - $\mathcal{X}_1 = \bar{\mathbf{x}} + (\sqrt{(1+\kappa)P_x}), \mathcal{W}_1 = \frac{1}{2(1+\kappa)}$,
  - $\mathcal{X}_2 = \bar{\mathbf{x}} - (\sqrt{(1+\kappa)P_x}), \mathcal{W}_2 = \frac{1}{2(1+\kappa)}$.
- Set of σ-points is transformed through the nonlinear function, i.e.
$$\mathcal{Y}_i = g(\mathcal{X}_i), \forall i.$$
- The desired characteristics are computed according to
  - $\bar{y}_A^{UKF} = \sum_{i=0}^{2} \mathcal{W}_i \mathcal{Y}_i$,
  - $P_{y,A}^{UKF} = \sum_{i=0}^{2} \mathcal{W}_i (\mathcal{Y}_i - \bar{y}_A^{UKF})(\mathcal{Y}_i - \bar{y}_A^{UKF})^T = P_{y,A}^{DD1} + P_{y,e}^{UKF}$,
  - $P_{xy,A}^{UKF} = \sum_{i=0}^{2} \mathcal{W}_i (\mathcal{X}_i - \bar{x})(\mathcal{Y}_i - \bar{y}_A^{UKF})^T = P_{xy,A}^{DD1}$.
- The accrual $P_{y,e}^{UKF}$ becomes significant especially for "highly" nonlinear function $g(\cdot)$ as well.

Common features of UKF and DDF's

Relation between Unscented Kalman and Divide Difference Filters

- Although the Unscented Kalman Filter arises from the totaly different idea then the Divide Difference Filters, it is possible to find their common features.
- Firstly, the relations of the UKF and the DD2 for the mean computation are the same, i.e.

$$\bar{\mathbf{y}}_A^{DD2} = \bar{\mathbf{y}}_A^{UKF},$$

  if the equality of the scaling parameters hold $n_x + \kappa = h^2$.
- The covariance matrixes obtained by the Stirling's interpolation 1st, 2nd order and by the unscented transformation are related as

$$\mathbf{P}_{y,A}^{DD1} \leq \mathbf{P}_{y,A}^{DD2} \approx \mathbf{P}_{y,A}^{UKF}.$$

- Moreover, for one-dimensional variables the UKF and the DD2 are the same filters if $n_x + \kappa = h^2$.

Numerical Example

Transformation of a Random Variable via Nonlinear Function

- The Gaussian random variable x is given by
  - the mean $\bar{x} = 2$,
  - the variance $P_x = 7$.
- The nonlinear function is in the form

$$y = g(x) = 0.2x^2 + 2.$$

	DD1 ($h^2 = 3$)	DD2 ($h^2 = 3$)	UKF ($\kappa = 2$)	MC
$\bar{y}$	2.80	4.20	4.20	4.20
P_y	4.48	8.40	8.40	8.40
$P_{y,e}$	–	3.92	3.92	–
P_{xy}	5.60	5.60	5.60	5.60

Impact of Chosen Approximation on Local Filter's Covariance Matrix

Local Filter Predictive Measurement Covariance Matrix

- The filtering covariance matrix $\mathbf{P}_k = cov[\mathbf{x}_k|\mathbf{z}^k]$ of the local filters can be computed as

$$\mathbf{P}_k = \mathbf{P}^{'}_k - \mathbf{P}^{'}_{xz,k}\mathbf{P}^{'-1}_{z,k}\mathbf{P}^{'T}_{xz,k},$$

  where $\mathbf{P}^{'}_k = \mathbf{P}^{'}_k(\mathbf{Q}_k)$ and $\mathbf{P}^{'}_{z,k} = \mathbf{P}^{'}_{z,k}(\mathbf{R}_k)$.
- Let us consider the significantly nonlinear function $\mathbf{h}_k(\cdot)$ in the measurement equation and the insignificant measurement noise covariance matrix $\mathbf{R}_k$ with respect to the state noise one $\mathbf{Q}_k$, then
  - $\mathbf{P}^{'DD1}_{z,k} \ll \mathbf{P}^{'DD2}_{z,k} \approx \mathbf{P}^{'UKF}_{z,k}$ and
  - $\mathbf{P}^{'DD1}_{xz,k} \approx \mathbf{P}^{'DD2}_{xz,k} \approx \mathbf{P}^{'UKF}_{xz,k}$

  follows from the properties of the particular approximation.
- The accruals $\mathbf{P}^{'UKF}_{z,e,k}$ and $\mathbf{P}^{'DD2}_{z,e,k}$ caused by the Stirling's interpolation 2nd order or the unscented transformation becomes significant.

Impact of Chosen Approximation on Local Filter's Covariance Matrix

Local Filter Filtering Covariance Matrix

- The properties of the covariance matrixes $\mathbf{P}'^{\cdots}_{z,k}$ and $\mathbf{P}'^{\cdots}_{xz,k}$ have the significant impact on the filtering covariance matrix, which can be expressed as follows

$$\mathbf{P}_k^{DD1} \ll \mathbf{P}_k^{DD2} \approx \mathbf{P}_k^{UKF}.$$

- In other words, it means that the DD1 can significantly reduce the filtering covariance matrix although the point estimate is not proportionally exact.
- However, the increase of $\mathbf{R}_k$ causes that the accruals $\mathbf{P}'^{UKF}_{z,e,k}$ and $\mathbf{P}'^{DD2}_{z,e,k}$ become insignificant and therefore
  - $\mathbf{P}'^{DD1}_{z,k} \approx \mathbf{P}'^{DD2}_{z,k} \approx \mathbf{P}'^{UKF}_{z,k}$,

  which leads to
  - $\mathbf{P}_k^{DD1} \approx \mathbf{P}_k^{DD2} \approx \mathbf{P}_k^{UKF}$.

Impact of Chosen Local Filter on Gaussian Sum Filter

Covariance Matrix of Gaussian Sum Filter

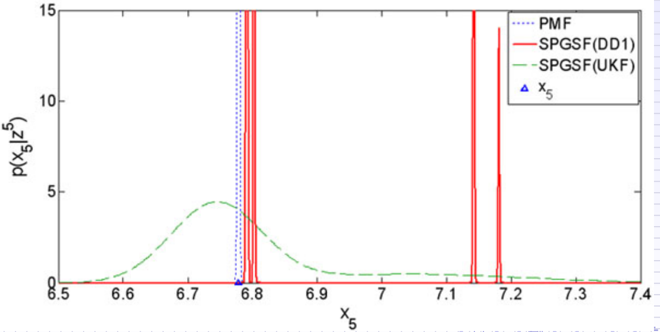
- The filtering covariance matrix of the Gaussian Sum Filter can be computed as

$$\mathbf{P}_k^{GSF} = \sum_{j=1}^{N_k} \beta_{j,k} \left(\mathbf{P}_{j,k}^{\cdots} + (\hat{\mathbf{x}}_k^{GSF} - \hat{\mathbf{x}}_{j,k}^{\cdots})(\hat{\mathbf{x}}_k^{GSF} - \hat{\mathbf{x}}_{j,k}^{\cdots})^T \right),$$

  where $\hat{\mathbf{x}}_k^{GSF} = \sum_{j=1}^{N_k} \beta_{j,k} \hat{\mathbf{x}}_{j,k}^{\cdots}$.
- The global filtering covariance matrix depends not only on the particular local covariance matrices $\mathbf{P}_{j,k}^{\cdots}$, but it is influenced by the placement of the local mean values $\hat{\mathbf{x}}_{j,k}^{\cdots}$ as well.

Quality of Filter's Point Estimates

	SPGSF(DD1)	SPGSF(UKF)	PMF
MSE	0.0222	0.0133	$3.55 \cdot 10^{-4}$
$\sum_{j=1}^{N} \mathbf{P}_{j,5}/N$	$1.77 \cdot 10^{-8}$	$8.59 \cdot 10^{-2}$	–
$\mathbf{P}_5^{GSF}$	$2.39 \cdot 10^{-2}$	$8.09 \cdot 10^{-2}$	$1.36 \cdot 10^{-6}$

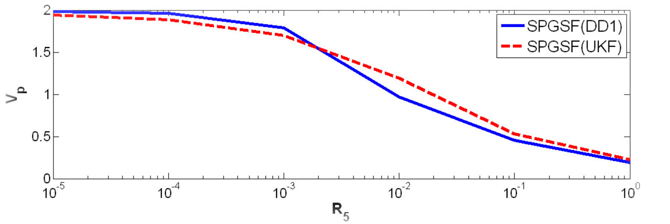


Dependence of Estimated PDF's on Noise Variance

- To compare the estimated pdf's of the Sigma Point Gaussian Sum Filters based on the Divide Difference Filters 1st Order or on the Unscented Kalman Filters the following criterion was set

$$V_p = \frac{\sum_{i=1}^{100} \int |p^i(x_5|z^5) - \hat{p}^i(x_5|z^5)|)}{100},$$

where $p^i(x_5|z^5)$ is the "true" pdf obtained by the PMF and $\hat{p}^i(x_5|z^5)$ is the pdf of the particular SPGSF.



Notes

- The particle filters exploiting the UKF (e.g. the Unscented Particle Filter, the Gaussian Mixture Sigma Point Particle Filter (van der Merwe and Wan, 2003)) suffer with similar shortcoming as the SPGSF(UKF).
- The "standard" GSF based on the bank of the Extended Kalman Filter shows similar properties to the SPGSF(DD1).
- The resultant estimates especially that of the DD2 and the UKF, and therefore the estimates of the SPGSF(DD2,UKF), heavily depend on the choice of the scaling parameters h^2 and κ, particularly when the system is highly nonlinear and the covariance matrix of the state noise is significantly different from that of the measurement noise.

Conclusion Remarks

- The derivative-free global and local filtering methods for the nonlinear systems were considered.
- It was shown that for some special cases the significant differences in the Sigma Point Gaussian Sum Filter's estimates could arise.
- The reasons of this unexpected behaviour were clarified by means of the analysis of the derivative-free local filters, namely the Divide Difference Filter 1st, 2nd Order and the Unscented Kalman Filter.
- During the analysis of the derivative-free local filters novel relations among these local filters were derived.
- It was shown that for one-dimensional systems the algorithms of the Divide Difference Filter 2nd Order and the Unscented Kalman Filter are the same.