Sigma Point Gaussian Sum Filter Design Using Square Root Unscented Filters

Miroslav Šimandl, Jindřich Duník

Department of Cybernetics
Faculty of Applied Sciences
University of West Bohemia in Pilsen
Czech Republic

7 July 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

The aim of the filtering is to find 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 Filtering Problem

Bayesian Recursive Relations (BRR's)

The solution of 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 BRR's

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

Solution of Filtering Problem

Bayesian Recursive Relations (BRR's)

The solution of 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 BRR's

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

Characteristic of Local Methods

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

Advantage and Disadvantage

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
- New derivative-free approach, e.g.
  Unscented Kalman Filter, Divide Difference Filter

Transformation of Random Variable

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

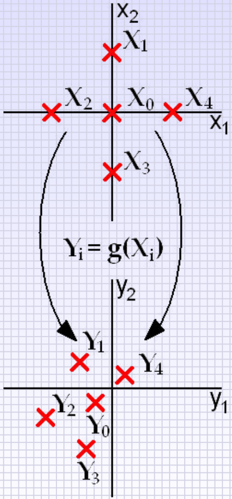
Unscented Transformation

- The random variable $\mathbf{x}$ is approximated by the set of deterministically chosen weighted points (so called σ-points). The σ-point set computation is given
  - $\mathcal{X}_0 = \bar{\mathbf{x}}, \mathcal{W}_0 = \frac{\kappa}{n_x+\kappa}$,
  - $\mathcal{X}_i = \bar{\mathbf{x}} + (\sqrt{(n_x+\kappa)\mathbf{P}_x})_i, i = 1, \ldots, n_x$,
  - $\mathcal{X}_j = \bar{\mathbf{x}} - (\sqrt{(n_x+\kappa)\mathbf{P}_x})_{j-n_x}, j = n_x+1, \ldots, 2n_x$,

  where $\mathcal{W}_i = \mathcal{W}_j = \frac{1}{2(n_x+\kappa)}, \forall i, j$.
- Set of σ-points are transformed through the nonlinear function, i.e.

$$\mathcal{Y}_i = \mathbf{g}(\mathcal{X}_i), \forall i.$$

- Desired characteristics are computed according to
  - $\bar{\mathbf{y}} = E[\mathbf{y}] \approx \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{Y}_i$,
  - $\mathbf{P}_y = E[(\mathbf{y}-\bar{\mathbf{y}})(\mathbf{y}-\bar{\mathbf{y}})^T] \approx \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{Y}_i - \bar{\mathbf{y}})(\mathcal{Y}_i - \bar{\mathbf{y}})^T$,
  - $\mathbf{P}_{xy} = E[(\mathbf{x}-\bar{\mathbf{x}})(\mathbf{y}-\bar{\mathbf{y}})^T] \approx \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{X}_i - \bar{\mathbf{x}})(\mathcal{Y}_i - \bar{\mathbf{y}})^T$.



Unscented Kalman Filter

- **Initialization:** Initial condition is assumed $p(\mathbf{x}_0|\mathbf{z}^{-1}) = \mathcal{N}\{\mathbf{x}_0 : \hat{\mathbf{x}}_0^{'}, \mathbf{P}_0^{'}\}$.
- Computation of predictive σ-point set $\{\mathcal{X}_{i,k|k-1}\}$, $\{\mathcal{W}_i\}$ from $\hat{\mathbf{x}}_k^{'}$, $\mathbf{P}_k^{'}$.
- **Filtering step:**
  - Predictive σ-points are transformed, i.e. $\mathcal{Z}_{i,k|k-1} = \mathbf{h}_k(\mathcal{X}_{i,k|k-1})$, $\forall i$.
  - The characteristics of the predictive measurement estimate, i.e. $\hat{\mathbf{z}}_k^{'}$, $\mathbf{P}_{z,k}^{'}$, and the cross-covariance matrix $\mathbf{P}_{xz,k}^{'}$ are calculated from the sets $\{\mathcal{X}_{i,k|k-1}\}$ and $\{\mathcal{Z}_{i,k|k-1}\}$.
  - The computation of the filtering mean and covariance matrix is performed:
    - $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^{'} + \mathbf{P}_{xz,k}^{'}\mathbf{P}_{z,k}^{'-1}(\mathbf{z}_k - \hat{\mathbf{z}}_k^{'})$,
    - $\mathbf{P}_k = \mathbf{P}_k^{'} - \mathbf{P}_{xz,k}^{'}\mathbf{P}_{z,k}^{'-1}\mathbf{P}_{xz,k}^{'T}$.
- Computation of filtering σ-point set $\{\mathcal{X}_{i,k|k}\}$ from $\hat{\mathbf{x}}_k$, $\mathbf{P}_k$.
- **Prediction step:**
  - Filtering σ-points are transformed, i.e. $\mathcal{X}_{i,k+1|k} = \mathbf{f}_k(\mathcal{X}_{i,k|k})$, $\forall i$.
  - The predictive mean $\hat{\mathbf{x}}_{k+1}^{'}$ and covariance matrix $\mathbf{P}_{k+1}^{'}$ is calculated from the set of σ-points $\{\mathcal{X}_{i,k+1|k}\}$.

Characteristic of Global Methods

The global methods are mainly based on an appropriate approximation of the description of the pdf's.

Advantage and Disadvantage

Advantage is

- certain convergence of the state estimate.

Disadvantage is

- growth of computational demands towards local methods.

Approaches in Global Estimation

- Analytical approach, e.g. Gaussian Sum Filter
- Numerical approach, e.g. Point-Mass Filter
- Simulation approach, e.g. Particle Filter

Disadvantages of Unscented Kalman Filer And Gaussian Sum Filter

- Unscented Kalman Filter (UKF)
  - The square root of state estimate covariance matrix is computed twice at each time instant for the σ-point set calculation.
  - The Cholesky decomposition is often used which is quite computational demanding and numerically unstable matrix operation.
- Gaussian Sum Filter (GSF)
  - For design of the GSF the derivations of nonlinear functions in the state and measurement equation are desired.

Goals of the Paper

- To derive a numerical stable version of the UKF (so called Square Root UKF (SRUKF)), where the square roots of state estimate covariance matrixes are directly available.
- To apply the SRUKF in the Gaussian sum framework to design a derivative-free GSF (so called Sigma Point GSF (SPGSF)).

Disadvantages of Unscented Kalman Filer And Gaussian Sum Filter

- Unscented Kalman Filter (UKF)
  - The square root of state estimate covariance matrix is computed twice at each time instant for the σ-point set calculation.
  - The Cholesky decomposition is often used which is quite computational demanding and numerically unstable matrix operation.
- Gaussian Sum Filter (GSF)
  - For design of the GSF the derivations of nonlinear functions in the state and measurement equation are desired.

Goals of the Paper

- To derive a numerical stable version of the UKF (so called Square Root UKF (SRUKF)), where the square roots of state estimate covariance matrixes are directly available.
- To apply the SRUKF in the Gaussian sum framework to design a derivative-free GSF (so called Sigma Point GSF (SPGSF)).

Design of Square Root Unscented Kalman Filter

- The computation of means remains without any change.
- The relations for computation of predictive covariance matrixes of the state $\mathbf{P}'_k$ and measurement $\mathbf{P}'_{z,k}$ can be easily transformed to square root form.
  - Consider the UKF relation for the predictive covariance matrix

$$\mathbf{P}'_k = \sum_{i=0}^{2n_x} \mathcal{W}_i(\mathcal{X}_{i,k|k-1} - \hat{\mathbf{x}}'_k)(\mathcal{X}_{i,k|k-1} - \hat{\mathbf{x}}'_k)^T + \mathbf{Q}_k.$$

  - It can be rewritten to the form $\mathbf{P}'_k = \mathbf{S}'_k \mathbf{S}'^T_k$, where

$$\mathbf{S}'_k = ht([\sqrt{\mathcal{W}_0}(\mathcal{X}_{0,k|k-1} - \hat{\mathbf{x}}'_k), \ldots, \sqrt{\mathcal{W}_{2n_x}}(\mathcal{X}_{2n_x,k|k-1} - \hat{\mathbf{x}}'_k), \mathbf{S}_{Q,k}]),$$

    $ht(\cdot)$ is Householder triangularization and $\mathbf{Q}_k = \mathbf{S}_{Q,k}\mathbf{S}^T_{Q,k}$ is state noise covariance matrix.
- The Householder triangularization can be applied to rectangular matrix $\mathbf{M}$ to obtain square matrix $\mathbf{N}$ so that the equality $\mathbf{M}\mathbf{M}^T = \mathbf{N}\mathbf{N}^T$ is fulfilled.

Design of Square Root Unscented Kalman Filter (cont'd)

- Transformation of the UKF relation for the filtering covariance matrix come out from the relation

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

- The square root form of the filtering covariance matrix can be expressed as

$$\mathbf{S}_k = ht([\mathbf{M}_{x,k}^{'} - \mathbf{K}_k \mathbf{M}_{z,k}^{'}, \mathbf{K}_k \mathbf{S}_{R,k}]),$$

  where $\mathbf{K}_k = \mathbf{M}_{x,k}^{'} \mathbf{M}_{z,k}^{'T} (\mathbf{S}_{z,k}^{'} \mathbf{S}_{z,k}^{'T})^{-1}$ is Kalman gain, $\mathbf{P}_{z,k}^{'} = \mathbf{S}_{z,k}^{'} \mathbf{S}_{z,k}^{'T}$ and $\mathbf{R}_k = \mathbf{S}_{R,k} \mathbf{S}_{R,k}^{T}$ is measurement noise covariance matrix.

- Rectangular matrixes $\mathbf{M}_{x,k}^{'}$ and $\mathbf{M}_{z,k}^{'T}$ consist of the predictive σ-points sets $\{\mathcal{X}_{i,k|k-1}\}$ and $\{\mathcal{Z}_{i,k|k-1}\}$, respectively.

Design of Sigma Point Gaussian Sum Filter

- **Initialization:** Initial condition is assumed

$$p(\mathbf{x}_0|\mathbf{z}^{-1}) = \sum_{j=1}^{N_0} w_k^{(j)} \mathcal{N}\{\mathbf{x}_0 : \hat{\mathbf{x}}_0'^{(j)}, \mathbf{S}_0'^{(j)}(\mathbf{S}_0'^{(j)})^T\}.$$

- **Filtering step:** Multiple application of the filtering part of the SRUKF for each pair $\hat{\mathbf{x}}_0'^{(j)}$ and $\mathbf{S}_0'^{(j)}$ leads to filtering pdf

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

- **Number of Gaussians reduction.**
- **Prediction step:** Multiple application of the prediction part of the SRUKF for each pair $\hat{\mathbf{x}}_k^{(j)}$ and $\mathbf{S}_k^{(j)}$ leads to prediction pdf

$$p(\mathbf{x}_{k+1}|\mathbf{z}^k) \approx \sum_{j=1}^{N_k} w_k^{(j)} \mathcal{N}\{\mathbf{x}_{k+1} : \hat{\mathbf{x}}_{k+1}'^{(j)}, \mathbf{S}_{k+1}'^{(j)}(\mathbf{S}_{k+1}'^{(j)})^T\}.$$

System Specification

Nonlinear Non-Gaussian System

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

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

where

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

Numerical Results

- The estimation performance of the Sigma Point Gaussian Sum Filter (**SPGSF**) is compared with the "standard" Gaussian Sum Filter (**GSF**), with the generic Particle Filter (**PF**) and with the Gaussian Mixture Sigma Point Particle Filter (**GMSPPF**) in the following table.

Algorithm	MSE	Time(s)
PF	1.9262	3.28
GMSPPF	0.0156	4.90
GSF	0.0253	0.91
SPGSF	0.0149	2.08

- The estimation performance of the Square Root Unscented Kalman Filter (**SRUKF**) is naturally the same as the "standard" Unscented Kalman Filter (**UKF**). However, the computational demands of the **SRUKF** are about 10% reduced towards the **UKF**.

Conclusion Remarks

- The problem of nonlinear derivative-free filters was considered.
- The square root version of the UKF, called Square Root Unscented Kalman Filter, was derived
  - to ensure the positive definiteness of covariance matrixes,
  - to reduce computational demands.
- The derivation technique of the SRUKF can be easily extended to the all versions of the UKF which differ in σ-point computation only.
- The SRUKF was put into the Gaussian sum framework to design the derivative-free global Sigma Point Gaussian Sum Filter.