Design of Derivative-Free Smoothers and Predictors

Miroslav Šimandl, Jindřich Duník

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

March 31, 2006

Description of System

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

where

- $\mathbf{f}_k(\cdot)$, $\mathbf{h}_k(\cdot)$ are known vector functions,
- $p_{\mathbf{w}_k}(\mathbf{w}_k) = \mathcal{N}\{\mathbf{w}_k : \mathbf{0}, \mathbf{Q}_k\}$, $p_{\mathbf{v}_k}(\mathbf{v}_k) = \mathcal{N}\{\mathbf{v}_k : \mathbf{0}, \mathbf{R}_k\}$, and $p_{\mathbf{x}_0}(\mathbf{x}_0)$ are known Gaussian probability density functions (pdf's).

State Estimation Problem

The aim of the state estimation is to find the probability density function of the state $\mathbf{x}_k$ conditioned by the measurements $\mathbf{z}^m = [\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_m]$.

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

Solution of the State Estimation Problem

Functional Recursive Relations (FRR's)

- $(k-m)$-step prediction $(k > m)$

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

- filtering $(m = k)$

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

- $(m-k)$-step smoothing $(m > k)$

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

Exact Solution of the FRR's

The closed-form solution of the FRR's is possible only for a few special cases, e.g. for linear Gaussian system.

Approximative Solution of the FRR's

One of the possible approximative solution of the FRR's is based on such approximation of the system description that the technique of the linear system estimators design can be used in the area of nonlinear systems as well.

Local Estimation Methods

The resultant estimates of the local estimators are given by the mean and the covariance matrix only.

Standard local estimation methods (1970)

The standard local estimation methods are based on the approximation of nonlinear functions in the state or the measurement equation by the Taylor expansion 1st or 2nd order

$$f_k(x_k) \approx f_k(\hat{x}_{k|k}) + \frac{df_k(x_k)}{dx_k}\Big|_{x_k=\hat{x}_{k|k}} \Delta x_k + \frac{d^2 f_k(x_k)}{dx_k^2}\Big|_{x_k=\hat{x}_{k|k}} (\Delta x_k)^2,$$

where $\Delta x_k = x_k - \hat{x}_{k|k}$.

These methods are used for multi-step prediction, filtering, and multi-step smoothing.

Novel local estimation methods (2000)

The novel local estimation methods are based on the Stirling's polynomial interpolation of the nonlinear functions or on the unscented transformation.

These methods are being developed mainly for filtering and one-step prediction.

Local Estimation Methods

- The main disadvantage of the standard local estimation methods can be found in the necessity of computation of the nonlinear functions derivatives.
- The novel estimation methods have been designed to preserve at least the same estimation performance of the standard ones without derivatives computation.
- However, the derivative-free multi-step smoothing and predictive methods have not been proposed yet.

Goal of the Paper

- To propose the multi-step derivative-free smoothers and predictors on the basis of the unscented transformation and the Stirling's polynomial interpolation.

Types of Smoothing

The smoothing problem $p(\mathbf{x}_k|\mathbf{z}^m)$, $k < m$ can be generally divided into three groups

- *fixed-point smoothing*, when k is fixed,
- *fixed-lag smoothing*, when difference $m - k$ is fixed, and
- *fixed-interval smoothing*, when time instant m is fixed.

Possible Solutions of Smoothing

The smoothing can be solved by two techniques based on

- the extension of the state $\mathbf{x}_m$ by the smoothed state $\mathbf{x}_k$, i.e.

$$\tilde{\mathbf{x}}_m = \begin{bmatrix} \mathbf{x}_m \\ \mathbf{x}_k \end{bmatrix},$$

  and application of common filtering algorithms,
- the solution of the particular FRR to propose a smoothing algorithm.

Rauch-Tung-Striebel Smoother and Linear Gaussian System

The exact solution of the smoothing problem $p(\mathbf{x}_k|z^m)$, $k < m$ for the linear Gaussian systems is given by the Rauch-Tung-Striebel smoother (RTSS)

$$\hat{\mathbf{x}}_{k|m} = E[\mathbf{x}_k|\mathbf{z}^m] = \hat{\mathbf{x}}_{k|k} + \mathbf{K}_{k|m}(\hat{\mathbf{x}}_{k+1|m} - \hat{\mathbf{x}}_{k+1|k}),$$
$$\mathbf{P}_{k|m} = cov[\mathbf{x}_k|\mathbf{z}^m] = \mathbf{P}_{k|k} - \mathbf{K}_{k|m}(\mathbf{P}_{k+1|k} - \mathbf{P}_{k+1|m})\mathbf{K}_{k|m}^T,$$

where $k = m-1, m-2, \ldots$,

- $\hat{\mathbf{x}}_{k+1|k}$, $\mathbf{P}_{k+1|k}$ are known predictive statistics,
- $\hat{\mathbf{x}}_{k|k}$, $\mathbf{P}_{k|k}$ are known filtering statistics,
- $\mathbf{K}_{k|m} = \mathbf{P}_{xx,k+1|k}(\mathbf{P}_{k+1|k})^{-1}$ is the smoothing gain, and
- $\mathbf{P}_{xx,k+1|k} = E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})^T|\mathbf{z}^k]$ is state cross-covariance matrix.

Rauch-Tung-Striebel Smoother and Linear Gaussian System (cont'd)

Linearity of the state equation ($\mathbf{f}_k(\mathbf{x}_k) = \mathbf{F}_k\mathbf{x}_k$) allows to find an exact solution of the state cross-covariance matrix in the form

$$\mathbf{P}_{xx,k+1|k} = E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})^T|\mathbf{z}^k] = \mathbf{P}_{k|k}\mathbf{F}_k^T.$$

Rauch-Tung-Striebel Smoother and Nonlinear Gaussian System

- The exact solution of the state cross-covariance matrix cannot be found for the system with the nonlinear state equation.
- The standard local methods approximate the nonlinear function $\mathbf{f}_k(\cdot)$ by the Taylor expansion which leads to

$$\mathbf{P}_{xx,k+1|k} = E[(\mathbf{x}_k-\hat{\mathbf{x}}_{k|k})(\mathbf{x}_{k+1}-\hat{\mathbf{x}}_{k+1|k})^T|\mathbf{z}^k] = \mathbf{P}_{k|k}\mathbf{F}_k^T(\hat{\mathbf{x}}_{k|k}),$$

where $\mathbf{F}_k(\hat{\mathbf{x}}_k) = \frac{\partial \mathbf{f}_k(\mathbf{x}_{k|k})}{\partial \mathbf{x}_k}|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k}}$.

Stirling's Interpolation 1st Order (Nørgaard, et al., 2000)

- The state nonlinear function $f_k(\cdot)$ can be approximated by means of the Stirling's Interpolation 1st Order, i.e.

$$x_{k+1} = f_k(x_k) + w_k \approx f_k(\hat{x}_{k|k}) + \frac{f_k(\hat{x}_{k|k} + \varepsilon) - f_k(\hat{x}_{k|k} - \varepsilon)}{2h} \Delta x_k + w_k,$$

  where $\varepsilon = h\sqrt{P_{k|k}}$ and $\Delta x_k = x_k - \hat{x}_{k|k}$.
- The predictive characteristics of the state x_{k+1} conditioned by the measurements z^k can be calculated in accord with
  - $\hat{x}_{k+1|k} = f_k(\hat{x}_{k|k})$,
  - $P_{k+1|k} = \frac{1}{4h^2}\left(f_k(\hat{x}_{k|k} + h\sqrt{P_{k|k}}) - f_k(\hat{x}_{k|k} - h\sqrt{P_{k|k}})\right)^2 + Q_k$,
  - $P_{xx,k+1|k} = \frac{\sqrt{P_{k|k}}}{2h}\left(f_k(\hat{x}_{k|k} + h\sqrt{P_{k|k}}) - f_k(\hat{x}_{k|k} - h\sqrt{P_{k|k}})\right)$.
- The Stirling's polynomial interpolation 2nd order can be employed as well.

Key Problem

To find an approximative solution of the state cross-covariance matrix

$$\mathbf{P}_{xx,k+1|k} = E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})^T|\mathbf{z}^k]$$

with the help of the Stirling's interpolation formula.

Approximative Solution of the State Cross-Covariance Matrix

The state cross-covariance matrix is given by

$$\mathbf{P}_{xx,k+1|k} = \mathbf{S}_{k|k}\mathbf{S}^T_{xx,k+1|k},$$

where

- $\mathbf{S}_{k|k}$ is a square-root of the filtering covariance matrix, i.e. $\mathbf{P}_{k|k} = \mathbf{S}_{k|k}\mathbf{S}^T_{k|k}$, and
- $\mathbf{S}_{xx,k+1|k} = \{S_{xx,k+1|k}(i,j)\} = \left\{\frac{f_{i,k}(\hat{\mathbf{x}}_{k|k}+h\mathbf{s}_{x,j})-f_{i,k}(\hat{\mathbf{x}}_{k|k}-h\mathbf{s}_{x,j})}{2h}\right\}$.

Unscented Transformation (Julier, et al., 2000)

- Another possibility is to approximate the pdf of the state estimate of x_k by a set of deterministically chosen weighted points (σ-points). This approximation technique is called the Unscented Transformation. The set is computed as
  - $\mathcal{X}^{(0)}_{k|k} = \hat{x}_{k|k}, \mathcal{W}_0 = \frac{\kappa}{1+\kappa}$,
  - $\mathcal{X}^{(1)}_{k|k} = \hat{x}_{k|k} + (\sqrt{(1+\kappa)P_{k|k}}), \mathcal{W}_1 = \frac{1}{2(1+\kappa)}$,
  - $\mathcal{X}^{(2)}_{k|k} = \hat{x}_{k|k} - (\sqrt{(1+\kappa)P_{k|k}}), \mathcal{W}_2 = \frac{1}{2(1+\kappa)}$.
- The set of σ-points is transformed through the nonlinear function, i.e.
$$\mathcal{X}^{(i)}_{k+1|k} = f_k(\mathcal{X}^{(i)}_{k|k}), \forall i.$$
- The desired characteristics are computed according to
  - $\hat{x}_{k+1|k} = \sum_{i=0}^{2} \mathcal{W}_i \mathcal{X}^{(i)}_{k+1|k}$,
  - $P_{k+1|k} = \sum_{i=0}^{2} \mathcal{W}_i (\mathcal{X}^{(i)}_{k+1|k} - \hat{x}_{k+1|k})(\mathcal{X}^{(i)}_{k+1|k} - \hat{x}_{k+1|k})^T + Q_k$,
  - $P_{xx,k+1|k} = \sum_{i=0}^{2} \mathcal{W}_i (\mathcal{X}^{(i)}_{k|k} - \hat{x}_{k|k})(\mathcal{X}^{(i)}_{k+1|k} - \hat{x}_{k+1|k})^T$.

Key Problem

To find an approximative solution of the cross-covariance matrix

$$\mathbf{P}_{xx,k+1|k} = E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k})^T|\mathbf{z}^k]$$

with the help of the unscented transformation.

Approximative Solution of the State Cross-Covariance Matrix

The state cross-covariance matrix is given by

$$\mathbf{P}_{xx,k+1|k} = \sum_{i=0}^{2n_x} \mathcal{W}_i(\mathcal{X}_{i,k|k} - \hat{\mathbf{x}}_{k|k})(\mathcal{X}^s_{i,k+1|k} - \hat{\mathbf{x}}^s_{k+1|k})^T,$$

where

- the σ-point set $\{\mathcal{W}_i, \mathcal{X}_{i,k|k}\}$ is computed from the filtering mean $\hat{\mathbf{x}}_{k|k}$ and the filtering covariance matrix $\mathbf{P}_{k|k}$,
- $\mathcal{X}^s_{i,k+1|k} = \mathbf{f}_k(\mathcal{X}_{i,k|k})$, $\forall i$, and
- the mean is $\hat{\mathbf{x}}^s_{k+1|k} = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{X}^s_{i,k+1|k}$.

Numerically stable version of Rauch-Tung-Striebel type smoothers

The square-root form of the smoothing covariance matrix

$$\mathbf{P}_{k|m} = \mathbf{S}_{k|m}\mathbf{S}_{k|m}^T = \mathbf{P}_{k|k} - \mathbf{K}_{k|m}(\mathbf{P}_{k+1|k} - \mathbf{P}_{k+1|m})\mathbf{K}_{k|m}^T,$$

can be expressed as

$$\mathbf{S}_{k|m} = ht\left([\mathbf{M}_{k|k} - \mathbf{K}_{k|m}\mathbf{M}_{k+1|k}, \mathbf{K}_{k|m}\mathbf{S}_{Q,k+1}, \mathbf{K}_{k|m}\mathbf{S}_{k+1|m}]\right),$$

where $ht(\cdot)$ is the Householder triangularization, $\mathbf{M}_{k|k}$ and $\mathbf{M}_{k+1|k}$ are known composed matrices, and $\mathbf{S}_{Q,k+1}$ is square-root of $\mathbf{Q}_{k+1}$.

Derivative-free smoothing algorithms

- Smoothing methods based on the extension of the state are suitable mainly for the *fixed-point smoothing*.
- The designed divided difference and unscented Rauch-Tung-Striebel smoothers can be easily applied for *all types of the smoothing problem*.

Multi-step local predictors

- The predictive pdf $p(\mathbf{x}_m|\mathbf{z}^k)$, $m > k$ is approximated by the Gaussian pdf. The predictive mean is given by

$$\begin{aligned}\hat{\mathbf{x}}_{m|k} = E[\mathbf{x}_m|\mathbf{z}^k] &= \int \mathbf{f}_{m-1}(\mathbf{x}_{m-1})p(\mathbf{x}_{m-1}|z^k)d\mathbf{x}_{m-1},\\ &\approx \sum_{i=0}^{2n_x} \mathcal{W}_i \mathbf{f}_{m-1}(\mathcal{X}_{i,m-1|k}),\\ &\approx \mathbf{f}_{m-1}(\hat{\mathbf{x}}_{m-1}), m = k+1, k+2, \ldots\end{aligned}$$

  for the unscented predictor and the 1st order divide difference predictor, respectively.
- Similarly the predictive covariance matrix can be computed.
- The multi-step prediction can be understood as a multiple application of the one-step prediction.

System Specification (Ito and Xiong, 2000)

$$x_{k+1} = x_k + 5\Delta t x_k(1 - x_k^2) + w_k,$$
$$z_k = \Delta t(x_k - 0.05)^2 + v_k,$$

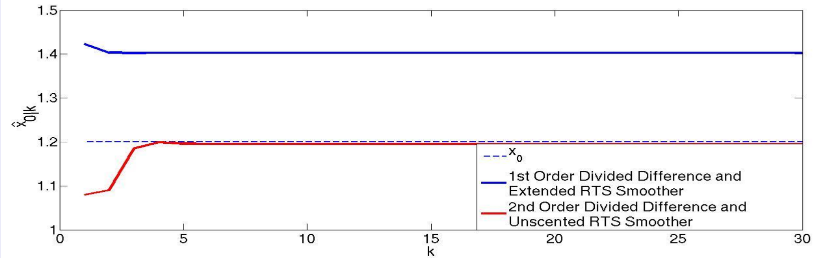
where

- $k = 1, \ldots, 400$, $\Delta t = 0.01$,
- $w_k \sim \mathcal{N}\{w_k : 0, 0.25\Delta t\}$, $v_k \sim \mathcal{N}\{v_k : 0, 0.01\Delta t\}$,
- system initial condition is $x_0 = 1.2$,
- and estimators initial condition is $p(x_0|z^{-1}) = \mathcal{N}\{x_0 : 2.2, 2\}$.

The aim is to estimate

- the initial condition $p(x_0|z^k)$ *(fixed-point smoothing)* and
- the smoothed pdf's $p(x_{k-10}|z^k)$ *(fixed-lag smoothing)*.

Estimate of the system initial condition *(fixed-point smoothing)*



Mean Square Error and Computational Time *(fixed-lag smoothing)*

Base of Algorithms	MSE of filtering	MSE of fixed-lag smoothing	smoothing time [s]
Tayl. exp. 1st ord.	0.1813	0.1491	0.0098
SI 1st ord.	0.1813	0.1491	0.0150
UT	0.1712	0.1310	0.0159

Conclusion

Conclusion Remarks

- The derivative-free estimation methods were discussed.
- The Stirling's interpolation and the unscented transformation were used for design of the derivative-free local unscented and divide difference smoothers and predictors.
- The numerical aspects of the Rauch-Tung-Striebel type smoothers were discussed and their numerical stable square-root versions were proposed.
- Now, the estimation problem (prediction, filtering, and smoothing) can be solved by the local derivative-free estimation methods.