MP-208

Optimal Filtering with Aerospace Applications

Chapter 7: Unscented Kalman Filter

Part I: Discrete-Time Formulation

Prof. Dr. Davi Antônio dos Santos
Instituto Tecnológico de Aeronáutica
www.professordavisantos.com

São José dos Campos - SP
2023

Contents

Problem Definition. . .

Problem Definition

State Equation:

Consider a state SP $\{\mathbf{X}_k\}$ and its realization $\{\mathbf{x}_k\}$, with $\mathbf{x}_k \in \mathbb{R}^{n_x}$ dynamically described by

$$\mathbf{x}_{k+1} = \mathbf{f}_k\left(\mathbf{x}_k, \mathbf{u}_k\right) + \mathbf{G}_k \mathbf{w}_k \tag{1}$$

where $\mathbf{u}_k \in \mathbb{R}^{n_u}$ is a known input, $\mathbf{w}_k \in \mathbb{R}^{n_w}$ is an unknown input, $\mathbf{f}_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_x}$ is a given non-linear function, and $\mathbf{G}_k \in \mathbb{R}^{n_x \times n_w}$ is a known matrix.

Assume that:

1. The initial state $\mathbf{x}_1$ is a realization of $\mathbf{X}_1$, which is assumed to be approximately symmetric and to have known mean $\bar{\mathbf{x}} \in \mathbb{R}^{n_x}$ and covariance $\bar{\mathbf{P}} \in \mathbb{R}^{n_x \times n_x}$. For short, we denote $\mathbf{X}_1 \sim (\bar{\mathbf{x}}, \bar{\mathbf{P}})$.
2. The sequence $\{\mathbf{w}_k\}$ is a realization of an uncorrelated SP $\{\mathbf{W}_k\}$, with an approx. symmetric $\mathbf{W}_k \sim (\mathbf{0}, \mathbf{Q}_k)$, where $\mathbf{Q}_k \in \mathbb{R}^{n_w \times n_w}$ is known.
3. $\left\{\{\mathbf{W}_k\}, \mathbf{X}_1\right\}$ is uncorrelated.

Problem Definition

Measurement Equation:

Consider a measurement SP $\{\mathbf{Y}_k\}$ and its realization $\{\mathbf{y}_k\}$, with $\mathbf{y}_{k+1} \in \mathbb{R}^{n_y}$ described by

$$\mathbf{y}_{k+1} = \mathbf{h}_{k+1}(\mathbf{x}_{k+1}) + \mathbf{v}_{k+1} \tag{2}$$

where $\mathbf{v}_{k+1} \in \mathbb{R}^{n_y}$ is an unknown input and $\mathbf{h}_{k+1} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}$ is a given non-linear function.

Assume that:

1. The sequence $\{\mathbf{v}_k\}$ is a realization of the uncorrelated SP $\{\mathbf{V}_k\}$, with approx. symmetric $\mathbf{V}_k \sim (\mathbf{0}, \mathbf{R}_k)$, where $\mathbf{R}_k \in \mathbb{R}^{n_y \times n_y}$ is known.
2. $\{\{\mathbf{V}_k\}, \{\mathbf{W}_k\}, \mathbf{X}_1\}$ is uncorrelated.

Problem Definition

Problem Statement:

The problem is to obtain an approximately optimal (MMSE) recursive filter for estimating $\{\mathbf{x}_k\}$ using $\{\mathbf{y}_k\}$, $\{\mathbf{u}_k\}$, and (1)–(2).

Comment:

In the previous section, we solved this problem using the EKF. Now, we formulate a different solution method: the unscented Kalman filter (UKF). Let us start by defining the so-called unscented transform (UT).

Unscented Transform…

Unscented Transform

Approximating the a Posteriori Distribution:

Consider a random vector $\mathbf{X} : \Omega \to \mathbb{R}^n$, $\mathbf{X} \sim (\bar{\mathbf{x}}, \mathbf{P}^x)$, an arbitrary function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ and the transformed RV $\mathbf{Y} : \Omega \to \mathbb{R}^m$, $\mathbf{Y} \sim (\bar{\mathbf{y}}, \mathbf{P}^y)$, obtained by

$$\mathbf{Y} = \mathbf{f}(\mathbf{X}) \tag{3}$$

The mean $\bar{\mathbf{y}}$ and the covariance $\mathbf{P}^y$ of $\mathbf{Y}$ can be approximated by the unscented transform (UT).

Unscented Transform

Procedure (UT):

1. Obtain the $2n+1$ σ-points $\mathcal{X}^i \in \mathbb{R}^n$, $i = 0, \ldots, 2n$, of $\mathbf{X}$, and the respective weights:

$$\mathcal{X}^0 = \bar{\mathbf{x}} \quad , \quad \rho^0 = \frac{\kappa}{n+\kappa} \tag{4}$$

$$\mathcal{X}^j = \bar{\mathbf{x}} + \sqrt{n+\kappa}\left(\sqrt{\mathbf{P}^x}\right)_j \quad , \quad \rho^j = \frac{1}{2(n+\kappa)} \tag{5}$$

$$\mathcal{X}^{j+n} = \bar{\mathbf{x}} - \sqrt{n+\kappa}\left(\sqrt{\mathbf{P}^x}\right)_j \quad , \quad \rho^{j+n} = \frac{1}{2(n+\kappa)} \tag{6}$$

   for $j = 1, \ldots, n$, where κ is a scale parameter; a common choice is $\kappa = 3 - n$.

2. Transform the σ-points $\mathcal{X}^i$ by $\mathbf{f}$, *i.e.*:

$$\mathcal{Y}^i = \mathbf{f}\left(\mathcal{X}^i\right) \in \mathbb{R}^m \tag{7}$$

   for $i = 0, \ldots, 2n$.

Unscented Transform

3. Approximate the *a posteriori* mean and covariance by sample statistics using the transformed σ-points $\mathcal{Y}^i$ and the weights defined in (4)–(7):

$$\bar{\mathbf{y}} \approx \sum_{i=0}^{2n} \rho^i \mathcal{Y}^i \tag{8}$$

$$\bar{\mathbf{P}}^y \approx \sum_{i=0}^{2n} \rho^i \left(\mathcal{Y}^i - \bar{\mathbf{y}}\right) \left(\mathcal{Y}^i - \bar{\mathbf{y}}\right)^{\mathrm{T}} \tag{9}$$

Unscented Transform

Comments:

1. Note that the set $\{\mathcal{X}^i, i = 0, ..., 2n\}$ is a deterministic sample of $\mathbf{X}$.
2. We know that, alternatively, the *a posteriori* mean and covariance can be approximated by first linearizing $\mathbf{f}$ and then using the linearity property of $E(.)$.
3. In general, the UT approximation is as good as the one obtained by the 2nd-order functional approximation of $\mathbf{f}$. In particular, if the *a priori* RV $\mathbf{X}$ is Gaussian, the UT is 3rd-order accurate.
4. The UT proposal was motivated by the fact that it is easier to approximate a probability distribution than a function (Julier & Uhlmann, 2004).
5. From now on, step 1 of the UT procedure is shortely denoted by

$$\left\{\left(\mathcal{X}^i, \rho^i\right),\ i = 0, ..., 2n\right\} \leftarrow \mathrm{SP}(\bar{\mathbf{x}}, \mathbf{P}^x) \tag{10}$$

Unscented Transform

Comments (cont.):

6. $\left(\sqrt{\mathbf{P}^x}\right)_j$ denotes the j-th column of $\sqrt{\mathbf{P}^x}$.
7. The sample mean and sample covariance of $\{\mathcal{X}^i, i = 0, ..., 2n\}$ are equal to the (theoretical) mean and covariance of $\mathbf{X}$, respectively.
8. Even though the weights ρ^i do not belong to the interval $[0, 1]$, their sum is equal to 1. In fact, this is a necessary condition for the property in item 7 to hold.

Discrete Unscented Kalman Filter...

Discrete Unscented Kalman Filter

Formulation Overview:

The Discrete Unscented Kalman Filter (DUKF) has the same structure as the DEKF. The only difference between them is that the former approximates the predictive expected values of the prediction phase by using the UT (instead of Taylor-series linearization).

Discrete Unscented Kalman Filter

Obtaining the σ-points:

Define the augmented state RV

$$ {}^a\mathbf{X}_k \triangleq \begin{bmatrix} \mathbf{X}_k \\ \mathbf{W}_k \\ \mathbf{V}_{k+1} \end{bmatrix} \in \mathbb{R}^{n_a} \tag{11} $$

where $n_a \triangleq n_x + n_w + n_y$.

From the problem definition and using the adopted notation for the filtered mean and covariance, the mean and covariance of ${}^a\mathbf{X}_k$ can be immediately obtained as

$$ {}^a\bar{\mathbf{x}}_k \triangleq \begin{bmatrix} \hat{\mathbf{x}}_{k|k} \\ \mathbf{0}_{n_w \times 1} \\ \mathbf{0}_{n_y \times 1} \end{bmatrix} \quad , \quad {}^a\mathbf{P}_k \triangleq \begin{bmatrix} \mathbf{P}_{k|k} & \mathbf{0}_{n_x \times n_w} & \mathbf{0}_{n_x \times n_y} \\ \mathbf{0}_{n_w \times n_x} & \mathbf{Q}_k & \mathbf{0}_{n_w \times n_y} \\ \mathbf{0}_{n_y \times n_x} & \mathbf{0}_{n_y \times n_w} & \mathbf{R}_{k+1} \end{bmatrix} \tag{12} $$

Discrete Unscented Kalman Filter

Obtaining the σ-points (cont.):

The σ-points ${}^{a}\mathcal{X}_k^i \in \mathbb{R}^{n_a}$ of the augmented state vector ${}^{a}\mathbf{X}_k$ are given by

$$\left\{ \left({}^{a}\mathcal{X}_k^i, \rho^i \right), i = 0, ..., 2n_a \right\} \leftarrow \mathrm{SP}\left({}^{a}\bar{\mathbf{x}}_k, {}^{a}\mathbf{P}_k \right) \tag{13}$$

and can be partitioned as

$${}^{a}\mathcal{X}_k^i \triangleq \begin{bmatrix} \mathcal{X}_k^i \\ \mathcal{W}_k^i \\ \mathcal{V}_{k+1}^i \end{bmatrix} \tag{14}$$

where $\mathcal{X}_k^i \in \mathbb{R}^{n_x}$, $\mathcal{W}_k^i \in \mathbb{R}^{n_w}$, and $\mathcal{V}_{k+1}^i \in \mathbb{R}^{n_y}$ are sample points of $\mathbf{X}_k$, $\mathbf{W}_k$, and $\mathbf{V}_{k+1}$, respectively.

Discrete Unscented Kalman Filter

Transforming the σ-points:

The σ-points $\mathcal{X}_k^i$ and $\mathcal{W}_k^i$, when transformed by (1), give rise to the σ-points of the predictive state, $\mathcal{X}_{k+1|k}^i \in \mathbb{R}^{n_x}$:

$$\mathcal{X}_{k+1|k}^i = \mathbf{f}_k\left(\mathcal{X}_k^i, \mathbf{u}_k\right) + \mathbf{G}_k \mathcal{W}_k^i \tag{15}$$

for $i = 0, ..., 2n_a$.

On the other hand, $\mathcal{X}_{k+1|k}^i$ and $\mathcal{V}_{k+1}^i$, when transformed by (2), give rise to the σ-points of the predictive measure:

$$\mathcal{Y}_{k+1|k}^i = \mathbf{h}_{k+1}\left(\mathcal{X}_{k+1|k}^i\right) + \mathcal{V}_{k+1}^i \tag{16}$$

for $i = 0, ..., 2n_a$.

Discrete Unscented Kalman Filter

Discrete-Time Prediction:

The predictive expected values are then immediately approximated by sample statistics on the predictive σ-points, *i.e.*:

$$\hat{\mathbf{x}}_{k+1|k} \approx \sum_{i=0}^{2n_a} \rho^i \; \mathcal{X}^i_{k+1|k} \tag{17}$$

$$\mathbf{P}_{k+1|k} \approx \sum_{i=0}^{2n_a} \rho^i \; \left(\mathcal{X}^i_{k+1|k} - \hat{\mathbf{x}}_{k+1|k}\right)\left(\mathcal{X}^i_{k+1|k} - \hat{\mathbf{x}}_{k+1|k}\right)^{\mathrm{T}} \tag{18}$$

$$\hat{\mathbf{y}}_{k+1|k} \approx \sum_{i=0}^{2n_a} \rho^i \; \mathcal{Y}^i_{k+1|k} \tag{19}$$

...

Discrete Unscented Kalman Filter

...

$$\mathbf{P}^{y}_{k+1|k} \approx \sum_{i=0}^{2n_a} \rho^i \left(\mathcal{Y}^i_{k+1|k} - \hat{\mathbf{y}}_{k+1|k}\right)\left(\mathcal{Y}^i_{k+1|k} - \hat{\mathbf{y}}_{k+1|k}\right)^{\mathrm{T}} \tag{20}$$

$$\mathbf{P}^{xy}_{k+1|k} \approx \sum_{i=0}^{2n_a} \rho^i \left(\mathcal{X}^i_{k+1|k} - \hat{\mathbf{x}}_{k+1|k}\right)\left(\mathcal{Y}^i_{k+1|k} - \hat{\mathbf{y}}_{k+1|k}\right)^{\mathrm{T}} \tag{21}$$

Discrete Unscented Kalman Filter

Update:

The update step of the DUKF is carried out as in the DEKF, *i.e.*, by computing

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}\left(\mathbf{y}_{k+1} - \hat{\mathbf{y}}_{k+1|k}\right) \tag{22}$$

$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{P}^{xy}_{k+1|k}\left(\mathbf{P}^{y}_{k+1|k}\right)^{-1}\left(\mathbf{P}^{xy}_{k+1|k}\right)^{\mathrm{T}} \tag{23}$$

where $\mathbf{K}_{k+1}$ is the Kalman gain, which is given by

$$\mathbf{K}_{k+1} = \mathbf{P}^{xy}_{k+1|k}\left(\mathbf{P}^{y}_{k+1|k}\right)^{-1} \tag{24}$$

Discrete Unscented Kalman Filter

Comments:

- Although the DUKF and DEKF have the same order of complexity (Wan & Merwe, 2000), in practice, the former is often more computationaly demanding. This is justified mostly by the need to compute a matrix square root (see equations (13) and (5)–(6)) at each filter iteration.
- For obtaining matrix square roots, we usually adopt the Cholesky factorization.
- Although in theory it is expected a better perfomance of the DUKF compared to the DEKF (for the same tuning), in problems that do not contain strong nonlinearities, they may be indistinguishable.

...

Discrete Uscented Kalman Filter

Comments (cont.):

- Note that the DUKF is easier to implement than the DEKF, since the former does not require the computation of Jacobians. Additionally, different from DEKF, the DUKF can be applied to systems containing non-smoothness in $\mathbf{f}_k$ or $\mathbf{h}_{k+1}$.
- Such as the EKF, for the UKF to be stable and to show a good convergence rate, we must tune the matrices $\bar{\mathbf{P}}$ and $\mathbf{Q}_k$ (we are going to see that in the next computational exercise).

References. . .

References

- Julier, S. J.; Uhlmann, J. K. Unscented Filtering and Nonlinear Estimation. **Proceedings of the IEEE**, 94(3), 2004.
- Wan, E. A.; Merwe, R. The Unscented Kalman Filter for Nonlinear Estimation. **IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium**, 2000.
- Sarkka, S. On Unscented Kalman Filtering for State Estimation of Continuous-Time Nonlinear Systems. **IEEE Transactions on Automatic Control**, 52(9), 2007.
- Wu, Y.; Hu, D.; Wu, M.; Hu, X.Unscented Kalman Filtering for Additive Noise Case: Augmented versus Nonaugmented. **IEEE Signal Precessing Letters**, 12(5), 2005.