ECE 286 Class 4: The Unscented Kalman Filter

Florian Meyer

Electrical and Computer Engineering Department
University of California San Diego

UC San Diego
JACOBS SCHOOL OF ENGINEERING

The Unscented Kalman Filter

- Consider a random vector $\mathbf{x} \in \mathbb{R}^L$ whose mean $\boldsymbol{\mu}_\mathbf{x}$ and covariance matrix $\mathbf{C}_\mathbf{x}$ are known, and a transformed random vector $\mathbf{y} = H(\mathbf{x})$

The Unscented Kalman Filter

- Consider a random vector $\mathbf{x} \in \mathbb{R}^L$ whose mean $\boldsymbol{\mu}_{\mathbf{x}}$ and covariance matrix $\mathbf{C}_{\mathbf{x}}$ are known, and a transformed random vector $\mathbf{y} = H(\mathbf{x})$
- Sigma points (SPs) are used to calculate approximations of the mean $\boldsymbol{\mu}_{\mathbf{y}}$, covariance matrix $\mathbf{C}_{\mathbf{y}}$, and cross-covariance matrix $\mathbf{C}_{\mathbf{x},\mathbf{y}}$

The Unscented Kalman Filter

- Consider a random vector $\mathbf{x} \in \mathbb{R}^L$ whose mean $\boldsymbol{\mu}_\mathbf{x}$ and covariance matrix $\mathbf{C}_\mathbf{x}$ are known, and a transformed random vector $\mathbf{y} = H(\mathbf{x})$
- Sigma points (SPs) are used to calculate approximations of the mean $\boldsymbol{\mu}_\mathbf{y}$, covariance matrix $\mathbf{C}_\mathbf{y}$, and cross-covariance matrix $\mathbf{C}_{\mathbf{x},\mathbf{y}}$
- The resulting approximations are at least as good as those obtained by linearizing $H(\cdot)$ as done in the extended Kalman filter
- Fundamental difference to the particle filter: The SPs are not random samples but are calculated by means of a deterministic algorithm

Calculation of Sigma Points

- For a L-dimensional random vector $\mathbf{x}$, SPs $\left\{\left(\mathbf{x}^{(j)}\right)\right\}_{j=1}^{2L}$ are calculated as

$$\mathbf{x}^{(j)} = \begin{cases} \boldsymbol{\mu}_{\mathbf{x}} + \sqrt{L}\left(\mathbf{C}_{\mathbf{x}}^{1/2}\right)_j, & j = 1, \ldots, L \\ \boldsymbol{\mu}_{\mathbf{x}} - \sqrt{L}\left(\mathbf{C}_{\mathbf{x}}^{1/2}\right)_j, & j = L+1, \ldots, 2L \end{cases}$$

Here, $\left(\mathbf{C}_{\mathbf{x}}^{1/2}\right)_j$ is the jth row or column of the matrix square root of $\mathbf{C}_{\mathbf{x}}$

S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, MIT Press, 2006

Calculation of Sigma Points

- For a L-dimensional random vector $\mathbf{x}$, SPs $\left\{(\mathbf{x}^{(j)})\right\}_{j=1}^{2L}$ are calculated as

$$\mathbf{x}^{(j)} = \begin{cases} \boldsymbol{\mu}_{\mathbf{x}} + \sqrt{L}\,\left(\mathbf{C}_{\mathbf{x}}^{1/2}\right)_j, & j = 1, \ldots, L \\ \boldsymbol{\mu}_{\mathbf{x}} - \sqrt{L}\,\left(\mathbf{C}_{\mathbf{x}}^{1/2}\right)_j, & j = L+1, \ldots, 2L \end{cases}$$

  Here, $\left(\mathbf{C}_{\mathbf{x}}^{1/2}\right)_j$ is the jth row or column of the matrix square root of $\mathbf{C}_{\mathbf{x}}$

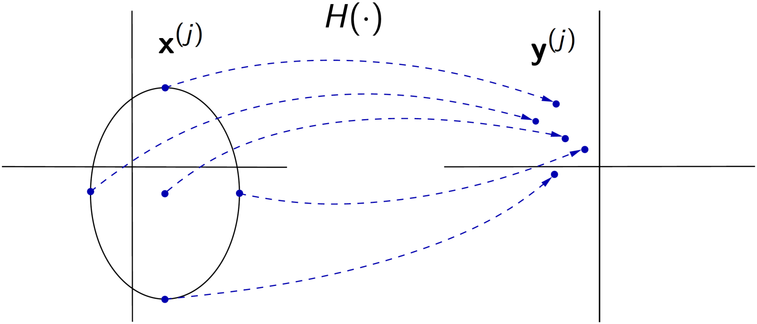
- SPs have the property that the sample mean $\tilde{\boldsymbol{\mu}}_{\mathbf{x}} = \frac{1}{2L}\sum_{j=1}^{2L}\mathbf{x}^{(j)}$ and sample covariance matrix $\tilde{\mathbf{C}}_{\mathbf{x}} = \frac{1}{2L}\sum_{j=1}^{2L}(\mathbf{x}^{(j)} - \tilde{\boldsymbol{\mu}}_{\mathbf{x}})(\mathbf{x}^{(j)} - \tilde{\boldsymbol{\mu}}_{\mathbf{x}})^{\mathrm{T}}$ are exactly equal to $\boldsymbol{\mu}_{\mathbf{x}}$ and $\mathbf{C}_{\mathbf{x}}$, respectively

S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, MIT Press, 2006

Unscented Transformation

- SPs $\{\mathbf{y}^{(j)}\}_{j=1}^{2L}$ of $\mathbf{y}$ can be obtained by propagating each SP $\mathbf{x}^{(j)}$ through $H(\cdot)$:

$$\mathbf{y}^{(j)} = H\big(\mathbf{x}^{(j)}\big), \qquad j = 1, \ldots, 2L$$



S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, MIT Press, 2006

Calculation of Mean and Covariance

- From $\{(\mathbf{x}^{(j)}, \mathbf{y}^{(j)}\}_{j=1}^{2L}$, one can calculate approximations of $\boldsymbol{\mu}_{\mathbf{y}}$, $\mathbf{C}_{\mathbf{y}}$ and $\mathbf{C}_{\mathbf{xy}}$ as

$$\tilde{\boldsymbol{\mu}}_{\mathbf{y}} = \frac{1}{2L} \sum_{j=1}^{2L} \mathbf{y}^{(j)}$$

$$\tilde{\mathbf{C}}_{\mathbf{y}} = \frac{1}{2L} \sum_{j=1}^{2L} (\mathbf{y}^{(j)} - \boldsymbol{\mu}_{\mathbf{y}})(\mathbf{y}^{(j)} - \boldsymbol{\mu}_{\mathbf{y}})^{\mathsf{T}}$$

$$\tilde{\mathbf{C}}_{\mathbf{xy}} = \frac{1}{2L} \sum_{j=1}^{2L} (\mathbf{y}^{(j)} - \boldsymbol{\mu}_{\mathbf{y}})(\mathbf{x}^{(j)} - \boldsymbol{\mu}_{\mathbf{x}})^{\mathsf{T}}$$

Bayesian Update

- SPs can be used for Bayesian estimation of a random vector $\mathbf{x}$ from an observed vector

  $$\mathbf{z} = \mathbf{y} + \mathbf{n} \quad \text{with } \mathbf{y} = H(\mathbf{x})$$

  Here, the noise $\mathbf{n}$ is zero-mean and statistically independent of $\mathbf{x}$, and has known covariance matrix $\mathbf{C_n}$

Bayesian Update

- SPs can be used for Bayesian estimation of a random vector $\mathbf{x}$ from an observed vector

  $$\mathbf{z} = \mathbf{y} + \mathbf{n} \quad \text{with } \mathbf{y} = H(\mathbf{x})$$

  Here, the noise $\mathbf{n}$ is zero-mean and statistically independent of $\mathbf{x}$, and has known covariance matrix $\mathbf{C_n}$

- Bayesian estimation relies on the posterior probability density function (pdf)

  $$f(\mathbf{x}|\mathbf{z}) \propto f(\mathbf{z}|\mathbf{x})f(\mathbf{x})$$

  where $f(\mathbf{z}|\mathbf{x})$ is the likelihood function and $f(\mathbf{x})$ is the prior pdf.

Bayesian Update

- SPs can be used for Bayesian estimation of a random vector $\mathbf{x}$ from an observed vector

$$\mathbf{z} = \mathbf{y} + \mathbf{n} \quad \text{with } \mathbf{y} = H(\mathbf{x})$$

  Here, the noise $\mathbf{n}$ is zero-mean and statistically independent of $\mathbf{x}$, and has known covariance matrix $\mathbf{C_n}$

- Bayesian estimation relies on the posterior probability density function (pdf)

$$f(\mathbf{x}|\mathbf{z}) \propto f(\mathbf{z}|\mathbf{x})f(\mathbf{x})$$

  where $f(\mathbf{z}|\mathbf{x})$ is the likelihood function and $f(\mathbf{x})$ is the prior pdf.

- Closed-form calculation of $f(\mathbf{x}|\mathbf{z})$ is usually infeasible

Bayesian Update in the Linear-Gaussian Case

- A feasible special case is when $H(\cdot)$ is linear, i.e., $H(\mathbf{x}) = \mathbf{H}\mathbf{x}$ and $\mathbf{x}$ and $\mathbf{n}$ are Gaussian. Then $f(\mathbf{x}|\mathbf{z})$ is also Gaussian, and $\boldsymbol{\mu}_{\mathbf{x}|\mathbf{z}}$ and $\mathbf{C}_{\mathbf{x}|\mathbf{z}}$ can be calculated as

$$\boldsymbol{\mu}_{\mathbf{x}|\mathbf{z}} = \boldsymbol{\mu}_{\mathbf{x}} + \mathbf{K}(\mathbf{z} - \boldsymbol{\mu}_{\mathbf{y}}) \qquad \mathbf{C}_{\mathbf{x}|\mathbf{z}} = \mathbf{C}_{\mathbf{x}} - \mathbf{K}(\mathbf{C}_{\mathbf{y}} + \mathbf{C}_{\mathbf{n}})\mathbf{K}^{\mathrm{T}}$$

with

$$\begin{aligned} \mathbf{K} &= \mathbf{C}_{\mathbf{xy}}(\mathbf{C}_{\mathbf{y}} + \mathbf{C}_{\mathbf{n}})^{-1} \\ \boldsymbol{\mu}_{\mathbf{y}} &= \mathbf{H}\boldsymbol{\mu}_{\mathbf{x}} \\ \mathbf{C}_{\mathbf{y}} &= \mathbf{H}\mathbf{C}_{\mathbf{x}}\mathbf{H}^{\mathrm{T}} \\ \mathbf{C}_{\mathbf{xy}} &= \mathbf{C}_{\mathbf{x}}\mathbf{H}^{\mathrm{T}} \end{aligned}$$

Bayesian Update with Sigma Points

- In the nonlinear case, $\boldsymbol{\mu}_{\mathbf{x}|\mathbf{z}}$ and $\mathbf{C}_{\mathbf{x}|\mathbf{z}}$ can be approximated by means of SPs

- This is done by using the closed-form expressions of the linear-Gaussian case, in which $\boldsymbol{\mu}_{\mathbf{y}}$, $\mathbf{C}_{\mathbf{y}}$, and $\mathbf{C}_{\mathbf{xy}}$ are replaced by the SP approximations $\tilde{\boldsymbol{\mu}}_{\mathbf{y}}$, $\tilde{\mathbf{C}}_{\mathbf{y}}$, and $\tilde{\mathbf{C}}_{\mathbf{xy}}$

Bayesian Update with Sigma Points

- In the nonlinear case, $\boldsymbol{\mu}_{\mathbf{x}|\mathbf{z}}$ and $\mathbf{C}_{\mathbf{x}|\mathbf{z}}$ can be approximated by means of SPs

- This is done by using the closed-form expressions of the linear-Gaussian case, in which $\boldsymbol{\mu}_{\mathbf{y}}$, $\mathbf{C}_{\mathbf{y}}$, and $\mathbf{C}_{\mathbf{xy}}$ are replaced by the SP approximations $\tilde{\boldsymbol{\mu}}_{\mathbf{y}}$, $\tilde{\mathbf{C}}_{\mathbf{y}}$, and $\tilde{\mathbf{C}}_{\mathbf{xy}}$

- SP approximations $\tilde{\boldsymbol{\mu}}_{\mathbf{x}|\mathbf{z}}$ and $\tilde{\mathbf{C}}_{\mathbf{x}|\mathbf{z}}$ are thus obtained as

$$\tilde{\boldsymbol{\mu}}_{\mathbf{x}|\mathbf{z}} = \boldsymbol{\mu}_{\mathbf{x}} + \tilde{\mathbf{K}}(\mathbf{z} - \tilde{\boldsymbol{\mu}}_{\mathbf{y}})\,, \qquad \tilde{\mathbf{C}}_{\mathbf{x}|\mathbf{z}} = \mathbf{C}_{\mathbf{x}} - \tilde{\mathbf{K}}(\tilde{\mathbf{C}}_{\mathbf{y}} + \mathbf{C}_{\mathbf{n}})\tilde{\mathbf{K}}^{\mathrm{T}},$$

with $\tilde{\mathbf{K}} = \tilde{\mathbf{C}}_{\mathbf{xy}}(\tilde{\mathbf{C}}_{\mathbf{y}} + \mathbf{C}_{\mathbf{n}})^{-1}$

Bayesian Update with Sigma Points

- In the nonlinear case, $\boldsymbol{\mu}_{\mathbf{x}|\mathbf{z}}$ and $\mathbf{C}_{\mathbf{x}|\mathbf{z}}$ can be approximated by means of SPs

- This is done by using the closed-form expressions of the linear-Gaussian case, in which $\boldsymbol{\mu}_{\mathbf{y}}$, $\mathbf{C}_{\mathbf{y}}$, and $\mathbf{C}_{\mathbf{xy}}$ are replaced by the SP approximations $\tilde{\boldsymbol{\mu}}_{\mathbf{y}}$, $\tilde{\mathbf{C}}_{\mathbf{y}}$, and $\tilde{\mathbf{C}}_{\mathbf{xy}}$

- SP approximations $\tilde{\boldsymbol{\mu}}_{\mathbf{x}|\mathbf{z}}$ and $\tilde{\mathbf{C}}_{\mathbf{x}|\mathbf{z}}$ are thus obtained as

$$\tilde{\boldsymbol{\mu}}_{\mathbf{x}|\mathbf{z}} = \boldsymbol{\mu}_{\mathbf{x}} + \tilde{\mathbf{K}}(\mathbf{z} - \tilde{\boldsymbol{\mu}}_{\mathbf{y}})\,, \qquad \tilde{\mathbf{C}}_{\mathbf{x}|\mathbf{z}} = \mathbf{C}_{\mathbf{x}} - \tilde{\mathbf{K}}(\tilde{\mathbf{C}}_{\mathbf{y}} + \mathbf{C}_{\mathbf{n}})\tilde{\mathbf{K}}^{\mathrm{T}},$$

  with $\tilde{\mathbf{K}} = \tilde{\mathbf{C}}_{\mathbf{xy}}(\tilde{\mathbf{C}}_{\mathbf{y}} + \mathbf{C}_{\mathbf{n}})^{-1}$

- Sequential implementation of this algorithm
  $\Rightarrow$ Unscented Kalman Filter