The Unscented Kalman Filter for State Estimation

Colin McManus

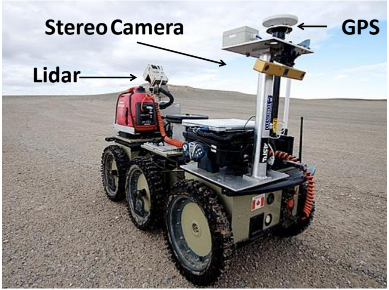
Autonomous Space Robotics Lab
University of Toronto Institute for Aerospace Studies

Presented at the Simultaneous Localization and Mapping (SLAM) Workshop

May 29th, 2010

Outline

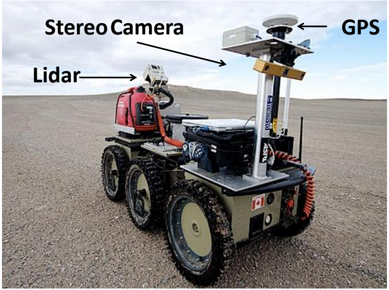
Defining terms



State:

$$p(\mathbf{x}_k | \mathbf{u}_{1:k}, \mathbf{y}_{1:k}) \rightarrow \mathcal{N}\left(\hat{\mathbf{x}}_k, \hat{\mathbf{P}}_k\right)$$

Defining terms



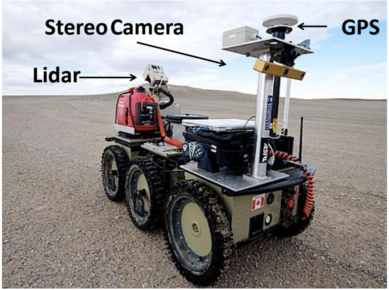
State:

$$p(\mathbf{x}_k | \mathbf{u}_{1:k}, \mathbf{y}_{1:k}) \rightarrow \mathcal{N}\left(\hat{\mathbf{x}}_k, \hat{\mathbf{P}}_k\right)$$

Motion model:

$$\mathbf{x}_k = \mathbf{h}\left(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{w}_k\right), \quad \mathbf{w}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{Q}_k\right)$$

Defining terms



State:

$$p(\mathbf{x}_k|\mathbf{u}_{1:k},\mathbf{y}_{1:k}) \rightarrow \mathcal{N}\left(\hat{\mathbf{x}}_k,\hat{\mathbf{P}}_k\right)$$

Motion model:

$$\mathbf{x}_k = \mathbf{h}\left(\mathbf{x}_{k-1},\mathbf{u}_k,\mathbf{w}_k\right), \quad \mathbf{w}_k \sim \mathcal{N}\left(\mathbf{0},\mathbf{Q}_k\right)$$

Sensor model:

$$\mathbf{y}_k = \mathbf{g}\left(\mathbf{x}_k,\mathbf{n}_k\right), \quad \mathbf{n}_k \sim \mathcal{N}\left(\mathbf{0},\mathbf{R}_k\right)$$

Goal

Goal at time-step k:

$$\{\hat{\mathbf{x}}_{k-1}, \hat{\mathbf{P}}_{k-1}, \mathbf{u}_k, \mathbf{y}_k\} \rightarrow \{\hat{\mathbf{x}}_k, \hat{\mathbf{P}}_k\}$$

Goal

Goal at time-step k:

$$\{\hat{\mathbf{x}}_{k-1}, \hat{\mathbf{P}}_{k-1}, \mathbf{u}_k, \mathbf{y}_k\} \rightarrow \{\hat{\mathbf{x}}_k, \hat{\mathbf{P}}_k\}$$

Measurement update (correction step):

$$\begin{aligned} \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_k^- + \mathbf{K}_k\left(\mathbf{y}_k - \hat{\mathbf{y}}_k\right) \\ \hat{\mathbf{P}}_k &= \hat{\mathbf{P}}_k^- - \mathbf{K}_k \mathbf{U}_k^T \end{aligned}$$

- $\hat{\mathbf{x}}_k^-$: Predicted state
- $\hat{\mathbf{P}}_k^-$: Predicted covariance
- $\mathbf{K}_k$: Kalman gain
- $\hat{\mathbf{y}}_k$: Predicted measurement
- $\mathbf{U}_k$: Cross-covariance term

Extended Kalman Filter

- Nonlinear extension to the famous 'Kalman Filter'
- EKF uses a first-order Taylor series expansion of the motion and observation models with respect to the current state estimate

Extended Kalman Filter

- Nonlinear extension to the famous 'Kalman Filter'
- EKF uses a first-order Taylor series expansion of the motion and observation models with respect to the current state estimate
- Prediction step:

$$\begin{aligned} \hat{\mathbf{x}}_k^- &= \mathbf{h}(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_k, \mathbf{0}) \\ \hat{\mathbf{P}}_k^- &= \mathbf{H}_{\mathbf{x},k} \hat{\mathbf{P}}_{k-1} \mathbf{H}_{\mathbf{x},k}^T + \mathbf{H}_{\mathbf{w},k} \mathbf{Q}_k \mathbf{H}_{\mathbf{w},k}^T \end{aligned}$$

where

$$\mathbf{H}_{\mathbf{x},k} := \left. \frac{\partial \mathbf{h}(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{w}_k)}{\partial \mathbf{x}_{k-1}} \right|_{\hat{\mathbf{x}}_{k-1}, \mathbf{u}_k, \mathbf{0}}, \quad \mathbf{H}_{\mathbf{w},k} := \left. \frac{\partial \mathbf{h}(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{w}_k)}{\partial \mathbf{w}_k} \right|_{\hat{\mathbf{x}}_{k-1}, \mathbf{u}_k, \mathbf{0}}.$$

Example

- Consider a simple example with a quadratic nonlinearity,

$$h\left(x_{k-1}, u_k, w_k\right) = x_{k-1}^2 + w_k$$

Example

(Loading)

Strengths and weaknesses

Strengths:

- Computationally inexpensive

Weaknesses:

- For highly nonlinear problems, the EKF is known to encounter issues with both accuracy and stability (Julier et al. ,1995; Wan and van der Merwe, 2000)
- When analytical Jacobians are not available, numerical Jacobians are required

Unscented Kalman Filter

- Introduced by Julier et al. (1995) as a derivative-free alternative to the EKF
- UKF uses a weighted set of deterministically sampled points called *sigma-points*, which are passed through the nonlinearity and are used to approximate the statistics of the distribution
- Unscented Transformation is accurate to at least third-order for Gaussian systems

Unscented Kalman Filter

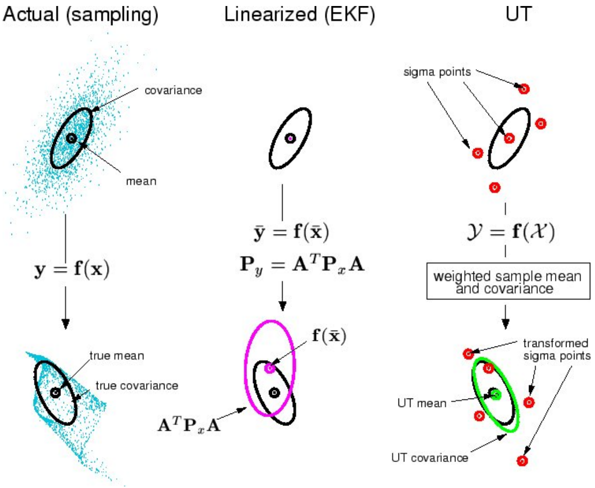


Image taken from van der Merwe and Wan (2001).

The Unscented Transformation

A set of $2N+1$ sigma-points is computed from the prior density, $\mathcal{N}\left(\hat{\mathbf{x}}_{k-1}, \hat{\mathbf{P}}_{k-1}\right)$, according to

$$\begin{aligned}
\mathbf{S}\mathbf{S}^T &:= \hat{\mathbf{P}}_{k-1} \quad \text{(Cholesky decomposition)} \\
\mathcal{X}_0 &:= \hat{\mathbf{x}}_{k-1} \\
\mathcal{X}_i &:= \hat{\mathbf{x}}_{k-1} + \sqrt{N+\kappa}\,\text{col}_i\mathbf{S} \\
\mathcal{X}_{i+N} &:= \hat{\mathbf{x}}_{k-1} - \sqrt{N+\kappa}\,\text{col}_i\mathbf{S}
\end{aligned} \qquad i = 1 \ldots N$$

where $N = \text{dim}(\hat{\mathbf{x}}_{k-1})$.

$$\mathcal{X}_i^- = \mathbf{h}\left(\mathcal{X}_i\right), \quad \text{for } i = 0 \ldots 2N$$

The Unscented Transformation

The mean and covariance are computed as follows:

$$\hat{\mathbf{x}}_k^- = \frac{1}{N+\kappa}\left(\kappa \mathcal{X}_0^- + \frac{1}{2}\sum_{i=1}^{2N} \mathcal{X}_i^-\right),$$

$$\hat{\mathbf{P}}_k^- = \frac{1}{N+\kappa}\left(\kappa\left(\mathcal{X}_0^- - \hat{\mathbf{x}}\right)\left(\mathcal{X}_0^- - \hat{\mathbf{x}}\right)^T + \frac{1}{2}\sum_{i=1}^{2N}\left(\mathcal{X}_i^- - \hat{\mathbf{x}}\right)\left(\mathcal{X}_i^- - \hat{\mathbf{x}}\right)^T\right).$$

The Unscented Transformation

The mean and covariance are computed as follows:

$$\hat{\mathbf{x}}_k^- = \frac{1}{N+\kappa}\left(\kappa\mathcal{X}_0^- + \frac{1}{2}\sum_{i=1}^{2N}\mathcal{X}_i^-\right),$$

$$\hat{\mathbf{P}}_k^- = \frac{1}{N+\kappa}\left(\kappa\left(\mathcal{X}_0^- - \hat{\mathbf{x}}\right)\left(\mathcal{X}_0^- - \hat{\mathbf{x}}\right)^T + \frac{1}{2}\sum_{i=1}^{2N}\left(\mathcal{X}_i^- - \hat{\mathbf{x}}\right)\left(\mathcal{X}_i^- - \hat{\mathbf{x}}\right)^T\right).$$

- Why this deterministic sampling scheme?

The Unscented Transformation

The mean and covariance are computed as follows:

$$\hat{\mathbf{x}}_k^- = \frac{1}{N+\kappa}\left(\kappa \mathcal{X}_0^- + \frac{1}{2}\sum_{i=1}^{2N}\mathcal{X}_i^-\right),$$

$$\hat{\mathbf{P}}_k^- = \frac{1}{N+\kappa}\left(\kappa\left(\mathcal{X}_0^- - \hat{\mathbf{x}}\right)\left(\mathcal{X}_0^- - \hat{\mathbf{x}}\right)^T + \frac{1}{2}\sum_{i=1}^{2N}\left(\mathcal{X}_i^- - \hat{\mathbf{x}}\right)\left(\mathcal{X}_i^- - \hat{\mathbf{x}}\right)^T\right).$$

- Why this deterministic sampling scheme?
- Consider the expectation of the prior mean plus a disturbance:

$$\hat{\mathbf{x}}_k^- = \mathsf{E}\left[\mathbf{h}\left(\hat{\mathbf{x}}_{k-1} + \Delta\mathbf{x}\right)\right], \quad \Delta\mathbf{x} \sim \mathcal{N}\left(\mathbf{0}, \hat{\mathbf{P}}_{k-1}\right)$$

Analytical Linearization

$$\begin{aligned}\hat{\mathbf{x}}_k^- &= \mathsf{E}\left[\mathbf{h}\left(\hat{\mathbf{x}}_{k-1} + \Delta\mathbf{x}\right)\right] \\ &= \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \mathsf{E}\left[\mathbf{D}_{\Delta\mathbf{x}}\mathbf{h} + \frac{\mathbf{D}^2_{\Delta\mathbf{x}}\mathbf{h}}{2!} + \frac{\mathbf{D}^3_{\Delta\mathbf{x}}\mathbf{h}}{3!} + \frac{\mathbf{D}^4_{\Delta\mathbf{x}}\mathbf{h}}{4!} + \dots\right]\end{aligned}$$

where,

$$\frac{\mathbf{D}^n_{\Delta\mathbf{x}}\mathbf{h}}{n!} = \frac{1}{n!}\left(\sum_{i=1}^{N} \Delta x_n \frac{\partial}{\partial x_n}\right)^n \mathbf{h}\left(\mathbf{x}\right)\Bigg|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}}$$

Analytical Linearization

By symmetry, odd-moments are zero,

$$\hat{\mathbf{x}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \mathsf{E}\left[\frac{\mathbf{D}^2_{\Delta\mathbf{x}}\mathbf{h}}{2!} + \frac{\mathbf{D}^4_{\Delta\mathbf{x}}\mathbf{h}}{4!} + \dots\right]$$

Analytical Linearization

By symmetry, odd-moments are zero,

$$\hat{\mathbf{x}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \mathsf{E}\left[\frac{\mathbf{D}_{\Delta\mathbf{x}}^2\mathbf{h}}{2!} + \frac{\mathbf{D}_{\Delta\mathbf{x}}^4\mathbf{h}}{4!} + \dots\right]$$

The second-order term is given by,

$$\mathsf{E}\left[\frac{\mathbf{D}_{\Delta\mathbf{x}}^2\mathbf{h}}{2!}\right] = \left(\frac{\nabla^T\mathsf{E}\left[\Delta\mathbf{x}\Delta\mathbf{x}^T\right]\nabla}{2!}\right)\mathbf{h}\Bigg|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}} = \left(\frac{\nabla^T\hat{\mathbf{P}}_{k-1}\nabla}{2!}\right)\mathbf{h}\Bigg|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}}$$

Analytical Linearization

By symmetry, odd-moments are zero,

$$\hat{\mathbf{x}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \mathsf{E}\left[\frac{\mathbf{D}^2_{\Delta\mathbf{x}}\mathbf{h}}{2!} + \frac{\mathbf{D}^4_{\Delta\mathbf{x}}\mathbf{h}}{4!} + \dots\right]$$

The second-order term is given by,

$$\mathsf{E}\left[\frac{\mathbf{D}^2_{\Delta\mathbf{x}}\mathbf{h}}{2!}\right] = \left(\frac{\nabla^T\mathsf{E}\left[\Delta\mathbf{x}\Delta\mathbf{x}^T\right]\nabla}{2!}\right)\mathbf{h}\Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}} = \left(\frac{\nabla^T\hat{\mathbf{P}}_{k-1}\nabla}{2!}\right)\mathbf{h}\Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}}$$

Resulting in the following

$$\hat{\mathbf{x}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \left(\frac{\nabla^T\hat{\mathbf{P}}_{k-1}\nabla}{2!}\right)\mathbf{h}\Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}} + \mathsf{E}\left[\frac{\mathbf{D}^4_{\Delta\mathbf{x}}\mathbf{h}}{4!} + \dots\right]$$

Statistical Linearization

We define some samples according to

$$\begin{aligned} \mathcal{X}_i &:= \hat{\mathbf{x}}_{k-1} + \sigma_i \\ \mathcal{X}_i^- &= \mathbf{h}(\mathcal{X}_i) \end{aligned} \qquad i = 0 \ldots 2N$$

where $N = \dim(\hat{\mathbf{x}}_{k-1})$. Now we calculate the weighted sigma-point expectations and compare with the 'true' mean

Statistical Linearization

$$\hat{\mathbf{x}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \frac{1}{2(N+\kappa)} \sum_{i=1}^{2N} \left(\mathbf{D}_{\sigma_i}\mathbf{h} + \frac{\mathbf{D}_{\sigma_i}^2\mathbf{h}}{2!} + \frac{\mathbf{D}_{\sigma_i}^3\mathbf{h}}{3!} + \frac{\mathbf{D}_{\sigma_i}^4\mathbf{h}}{4!} + \dots \right)$$

Statistical Linearization

$$\hat{\mathbf{x}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \frac{1}{2(N+\kappa)} \sum_{i=1}^{2N} \left(\mathbf{D}_{\sigma_i}\mathbf{h} + \frac{\mathbf{D}_{\sigma_i}^2\mathbf{h}}{2!} + \frac{\mathbf{D}_{\sigma_i}^3\mathbf{h}}{3!} + \frac{\mathbf{D}_{\sigma_i}^4\mathbf{h}}{4!} + \dots \right)$$

The second-order term is given by

$$\frac{\mathbf{D}_{\sigma_i}^2\mathbf{h}}{2!} = \left(\frac{\nabla^T \sigma_i \sigma_i^T \nabla}{2!} \right) \mathbf{h} \Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}}$$

Statistical Linearization

$$\hat{\mathbf{x}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \frac{1}{2(N+\kappa)}\sum_{i=1}^{2N}\left(\mathbf{D}_{\sigma_i}\mathbf{h} + \frac{\mathbf{D}_{\sigma_i}^2\mathbf{h}}{2!} + \frac{\mathbf{D}_{\sigma_i}^3\mathbf{h}}{3!} + \frac{\mathbf{D}_{\sigma_i}^4\mathbf{h}}{4!} + \dots\right)$$

The second-order term is given by

$$\frac{\mathbf{D}_{\sigma_i}^2\mathbf{h}}{2!} = \left(\frac{\nabla^T\sigma_i\sigma_i^T\nabla}{2!}\right)\mathbf{h}\Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}}$$

Recall

$$\mathsf{E}\left[\frac{\mathbf{D}_{\Delta\mathbf{x}}^2\mathbf{h}}{2!}\right] = \left(\frac{\nabla^T\mathsf{E}\left[\Delta\mathbf{x}\Delta\mathbf{x}^T\right]\nabla}{2!}\right)\mathbf{h}\Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}} = \left(\frac{\nabla^T\hat{\mathbf{P}}_{k-1}\nabla}{2!}\right)\mathbf{h}\Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}}$$

Statistical Linearization

We let $\sigma_i := \pm\sqrt{N+\kappa}\text{col}_i\mathbf{S}$, where $\mathbf{S}\mathbf{S}^T = \hat{\mathbf{P}}_{k-1}$, which gives

$$\hat{\mathbf{x}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \left(\frac{\nabla^T\hat{\mathbf{P}}_{k-1}\nabla}{2!}\right)\mathbf{h}\Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}} + \frac{1}{2(N+\kappa)}\sum_{i=1}^{2N}\left(\frac{\mathbf{D}_{\sigma_i}^4\mathbf{h}}{4!} + \dots\right)$$

Statistical Linearization

We let $\sigma_i := \pm\sqrt{N+\kappa}\,\text{col}_i\mathbf{S}$, where $\mathbf{S}\mathbf{S}^T = \hat{\mathbf{P}}_{k-1}$, which gives

$$\hat{\mathbf{x}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \left(\frac{\nabla^T \hat{\mathbf{P}}_{k-1} \nabla}{2!}\right) \mathbf{h}\bigg|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}} + \frac{1}{2(N+\kappa)} \sum_{i=1}^{2N} \left(\frac{\mathbf{D}_{\sigma_i}^4 \mathbf{h}}{4!} + \dots \right)$$

Compare above with analytical linearization

$$\hat{\mathbf{x}}_k^- = \mathbf{h}\left(\hat{\mathbf{x}}_{k-1}\right) + \left(\frac{\nabla^T \hat{\mathbf{P}}_{k-1} \nabla}{2!}\right) \mathbf{h}\bigg|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}} + \mathsf{E}\left[\frac{\mathbf{D}_{\Delta\mathbf{x}}^4 \mathbf{h}}{4!} + \dots \right]$$

Example

(Loading)

Augmented UKF

$$\mathbf{x}_k = \mathbf{h}\left(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{w}_k\right), \qquad \mathbf{w}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{Q}_k\right)$$
$$\mathbf{y}_k = \mathbf{g}\left(\mathbf{x}_k, \mathbf{n}_k\right), \qquad \mathbf{n}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{R}_k\right)$$

Augmented UKF

$$\begin{aligned} \mathbf{x}_k &= \mathbf{h}\left(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{w}_k\right), \qquad & \mathbf{w}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{Q}_k\right) \\ \mathbf{y}_k &= \mathbf{g}\left(\mathbf{x}_k, \mathbf{n}_k\right), \qquad & \mathbf{n}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{R}_k\right) \end{aligned}$$

- Completely general case, both the prior belief, the process noise and measurement noise have uncertainty so these are stacked together in the following way:

$$\mathbf{z} := \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \qquad \mathbf{Y} := \begin{bmatrix} \hat{\mathbf{P}}_{k-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_k \end{bmatrix}$$

Augmented UKF

$$\begin{aligned} \mathbf{x}_k &= \mathbf{h}\left(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{w}_k\right), & \mathbf{w}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{Q}_k\right) \\ \mathbf{y}_k &= \mathbf{g}\left(\mathbf{x}_k, \mathbf{n}_k\right), & \mathbf{n}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{R}_k\right) \end{aligned}$$

- Completely general case, both the prior belief, the process noise and measurement noise have uncertainty so these are stacked together in the following way:

$$\mathbf{z} := \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \qquad \mathbf{Y} := \begin{bmatrix} \hat{\mathbf{P}}_{k-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_k \end{bmatrix}$$

- Let $N := \dim\left(\hat{\mathbf{x}}\right)$, $P := \dim\left(\mathbf{Q}_k\right)$, $M := \dim\left(\mathbf{R}_k\right)$
- Number of sigma-points: $2(N + P + M) + 1$

Additive noise

- If our motion and observation model are given by

$$\begin{aligned} \mathbf{x}_k &= \mathbf{h}\left(\mathbf{x}_{k-1}, \mathbf{u}_k\right) + \mathbf{w}_k, \quad & \mathbf{w}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{Q}_k\right) \\ \mathbf{y}_k &= \mathbf{g}\left(\mathbf{x}_k\right) + \mathbf{n}_k, & \mathbf{n}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{R}_k\right) \end{aligned}$$

- Only $2N+1$ sigma-points are required!

$$\mathbf{z} := \hat{\mathbf{x}}_{k-1}, \qquad \mathbf{Y} := \hat{\mathbf{P}}_{k-1}$$

- Must re-draw sigma-points from $\{\hat{\mathbf{x}}_k^-, \hat{\mathbf{P}}_k^-\}$ for the correction step (lose odd-moment information)
- Alternatively, we can use $2(N+P)+1$ sigma-points for the correction step and avoid re-drawing (incorporate $\sqrt{\mathbf{Q}_k}$ into sigma-points)

Additive noise

- If our motion and observation model are given by

$$\begin{aligned} \mathbf{x}_k &= \mathbf{h}\left(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{w}_k\right), & \mathbf{w}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{Q}_k\right) \\ \mathbf{y}_k &= \mathbf{g}\left(\mathbf{x}_k\right) + \mathbf{n}_k, & \mathbf{n}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{R}_k\right) \end{aligned}$$

- Only $2(N+P)+1$ sigma-points are required

$$\mathbf{z} := \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \mathbf{0} \end{bmatrix}, \qquad \mathbf{Y} := \begin{bmatrix} \hat{\mathbf{P}}_{k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_k \end{bmatrix}$$

Other efficiency improvements

Other efficiency improvements

- Spherical-simplex sigma-points (Julier et al., 2003)
  - Used a set of $N + 2$ sigma-points, which were chosen to minimize the third-order moments (accurate to second-order)

Other efficiency improvements

- Spherical-simplex sigma-points (Julier et al., 2003)
  - Used a set of $N + 2$ sigma-points, which were chosen to minimize the third-order moments (accurate to second-order)
- Reduced sigma-point UKF (Quine, 2006)
  - Used a minimal set of $N + 1$ sigma-points (accurate to second-order)

Other efficiency improvements

- Spherical-simplex sigma-points (Julier et al., 2003)
  - Used a set of $N + 2$ sigma-points, which were chosen to minimize the third-order moments (accurate to second-order)
- Reduced sigma-point UKF (Quine, 2006)
  - Used a minimal set of $N + 1$ sigma-points (accurate to second-order)
- Square-root Unscented Kalman Filter (van der Merwe and Wan, 2001)
  - Efficient square-root form that avoids refactorizing the state covariance at the prediction step and reduces computational cost required to compute the Kalman gain
  - Present additional benefits in terms of numerical stability and ensures that the state covariance is always positive-definite

Strengths and weaknesses

Strengths:

- Unscented Transformation is accurate to third-order for Gaussian systems
- Derivative-free (easy implementation)

Weaknesses:

- Depending on noise assumptions, can be expensive
- Sigma-point scaling issues

Strengths and weaknesses

Strengths:

- Unscented Transformation is accurate to third-order for Gaussian systems
- Derivative-free (easy implementation)

Weaknesses:

- Depending on noise assumptions, can be expensive
- Sigma-point scaling issues
  - The Scaled Unscented Kalman Filter (Julier, 2002)

Strengths and weaknesses

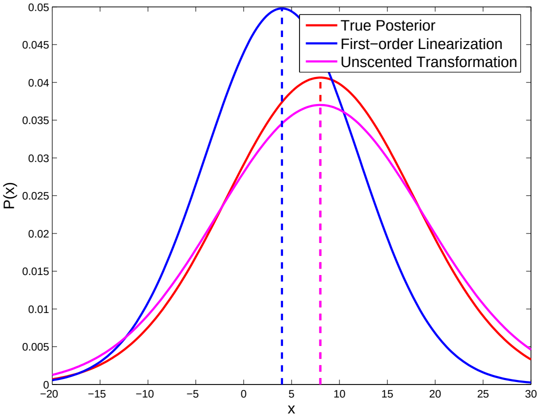
Strengths:

- Unscented Transformation is accurate to third-order for Gaussian systems
- Derivative-free (easy implementation)

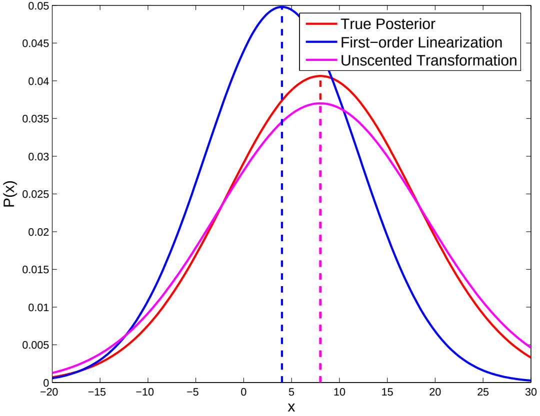
Weaknesses:

- Depending on noise assumptions, can be expensive
- Sigma-point scaling issues
  - The Scaled Unscented Kalman Filter (Julier, 2002)
  - Additional weight parameters (Wan and van der Merwe, 2000)

Accuracy


True Posterior
First−order Linearization
Unscented Transformation
P(x)
X
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
−20
−15
−10
−5
0
5
10
15
20
25
30

Accuracy



- Tong, C. and Barfoot, T.D. “A Comparison of the EKF, SPKF, and the Bayes Filter for Landmark-Based Localization,” In Proceedings of the 7th Canadian Conference on Computer and Robot Vision (CRV), to appear. Ottawa, Canada, 31 May - 2 June 2010.

Computational cost

- In general, EKF cost $<$ UKF cost

Computational cost

- In general, EKF cost $<$ UKF cost
- Problem dependent
- UKF form:
  - Additive UKF
  - Augmented UKF

Computational cost

- In general, EKF cost $<$ UKF cost
- Problem dependent
- UKF form:
  - Additive UKF
  - Augmented UKF
- Sigma point reduction:
  - Spherical-simplex sigma-points (Julier et al., 2003)
  - Reduced sigma-point UKF (Quine, 2006)

Any Questions?

References

- Julier, S., Uhlmann, J. and Durrant-Whyte, H. “A New Approach for Filtering Nonlinear Systems,” Proceedings of the American Control Conference, 1995, 3, 1628-1632.
- Julier, S. “The Scaled Unscented Transform,” Proceedings of the American Control Conference, 2002, 6, 4555 - 4559.
- Julier, S. “The Spherical Simplex Unscented Transformation,” Proceedings of the American Control Conference, 2003.
- Quine, B. “A derivative-free implementation of the extended Kalman filter,” Automatica, 2006, 42, 19271934.

References

- Tong, C. and Barfoot, T.D. “A Comparison of the EKF, SPKF, and the Bayes Filter for Landmark-Based Localization,” In Proceedings of the 7th Canadian Conference on Computer and Robot Vision (CRV), to appear. Ottawa, Canada, 31 May - 2 June 2010.
- van der Merwe, R. and Wan, E. “The Square-Root Unscented Kalman Filter for State and Parameter-Estimation,” IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001, 6, 3461-3464.
- Wan, E. and van der Merwe, R. “The Unscented Kalman Filter for Nonlinear Estimation,” Adaptive Systems for Signal Processing, Communications, and Control Symposium, 2000, 153-158.