ECE276A: Sensing & Estimation in Robotics
Lecture 11: Extended and Unscented Kalman Filtering

Instructor:

Nikolay Atanasov: natanasov@ucsd.edu

Teaching Assistants:

Qiaojun Feng: qif007@eng.ucsd.edu

Arash Asgharivaskasi: aasghari@eng.ucsd.edu

Thai Duong: tduong@eng.ucsd.edu

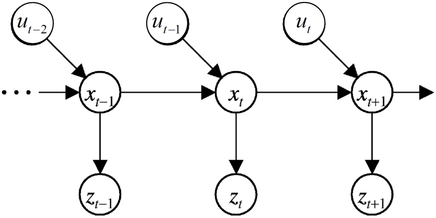
Yiran Xu: y5xu@eng.ucsd.edu

UC San Diego

JACOBS SCHOOL OF ENGINEERING

Electrical and Computer Engineering

Bayes Filter



- **Motion model**:
  $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) \sim p_f(\cdot \mid \mathbf{x}_t, \mathbf{u}_t)$
- **Observation model**:
  $\mathbf{z}_t = h(\mathbf{x}_t, \mathbf{v}_t) \sim p_h(\cdot \mid \mathbf{x}_t)$
- **Filtering**: keeps track of

$$p_{t|t}(\mathbf{x}_t) := p(\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1})$$
$$p_{t+1|t}(\mathbf{x}_{t+1}) := p(\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t})$$

- **Bayes filter**:

$$p_{t+1|t+1}(\mathbf{x}_{t+1}) = \underbrace{\overbrace{\frac{1}{p(\mathbf{z}_{t+1}|\mathbf{z}_{0:t}, \mathbf{u}_{0:t})}}^{\frac{1}{\eta_{t+1}}} p_h(\mathbf{z}_{t+1} \mid \mathbf{x}_{t+1}) \overbrace{\int p_f(\mathbf{x}_{t+1} \mid \mathbf{x}_t, \mathbf{u}_t) p_{t|t}(\mathbf{x}_t) d\mathbf{x}_t}^{\textbf{Predict: } p_{t+1|t}(\mathbf{x}_{t+1})}}_{\textbf{Update}}$$

Kalman Filter

- The **Kalman filter** is a Bayes filter with the following **assumptions**:
  - The prior pdf $p_{0|0}$ is Gaussian
  - The motion model is linear in the state and affected by Gaussian noise
  - The observation model is linear in the state and affected by Gaussian noise
  - The process noise $\mathbf{w}_t$ and measurement noise $\mathbf{v}_t$ are independent of each other, of the state $\mathbf{x}_t$, and across time

- **Prior**: $\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \Sigma_{t|t})$

- **Motion Model**:

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) := F\mathbf{x}_t + G\mathbf{u}_t + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(0, W)$$

$$\mathbf{x}_{t+1} \mid \mathbf{x}_t, \mathbf{u}_t \sim \mathcal{N}(F\mathbf{x}_t + G\mathbf{u}_t, W), \quad F \in \mathbb{R}^{d_x \times d_x}, G \in \mathbb{R}^{d_x \times d_u}, \; W \in \mathbb{R}^{d_x \times d_x}$$

- **Observation Model**:

$$\mathbf{z}_t = h(\mathbf{x}_t, \mathbf{v}_t) := H\mathbf{x}_t + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathcal{N}(0, V)$$

$$\mathbf{z}_t \mid \mathbf{x}_t \sim \mathcal{N}(H\mathbf{x}_t, V), \quad H \in \mathbb{R}^{d_z \times d_x}, \; V \in \mathbb{R}^{d_z \times d_z}$$

Nonlinear Kalman Filter

- A **nonlinear Kalman filter** is a Bayes filter such that:
  - The prior pdf $p_{0|0}$ is Gaussian
  - The motion model is ~~linear in the state and~~ affected by Gaussian noise
  - The observation model is ~~linear in the state and~~ affected by Gaussian noise
  - The process noise $\mathbf{w}_t$ and measurement noise $\mathbf{v}_t$ are independent of each other, of the state $\mathbf{x}_t$ and across time
  - The posterior pdf is **forced to be Gaussian via approximation**

- **Prior**: $\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \Sigma_{t|t})$

- **Motion Model**: $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \quad \mathbf{w}_t \sim \mathcal{N}(0, W)$

- **Observation Model**: $\mathbf{z}_t = h(\mathbf{x}_t, \mathbf{v}_t), \quad \mathbf{v}_t \sim \mathcal{N}(0, V)$

- **Challenge**: the predicted and updated pdfs are not Gaussian and can no longer be evaluated in closed form

- **Moment matching**: we can force the predicted and updated pdfs to be Gaussian by evaluating their first and second moments and approximating them with Gaussians with the same moments

Moment Matching

- Let $\mathbf{y} = f(\mathbf{x})$ be a nonlinear transformation of $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$
- The mean and (co)variance of $\mathbf{y}$ are:

$$\mathbf{m} := \mathbb{E}[\mathbf{y}] = \int f(\mathbf{x})\phi(\mathbf{x}; \boldsymbol{\mu}, \Sigma) d\mathbf{x}$$

$$S := \mathbb{E}\left[(\mathbf{y} - \mathbb{E}[\mathbf{y}])(\mathbf{y} - \mathbb{E}[\mathbf{y}])^\top\right] = \mathbb{E}\left[\mathbf{y}\mathbf{y}^\top\right] - \mathbb{E}[\mathbf{y}]\mathbb{E}[\mathbf{y}]^\top$$

$$= \int f(\mathbf{x}) f(\mathbf{x})^\top \phi(\mathbf{x}; \boldsymbol{\mu}, \Sigma) d\mathbf{x} - \mathbf{m}\mathbf{m}^\top$$

- The covariance of $\mathbf{x}$ and $\mathbf{y}$ is:

$$C := \mathbb{E}\left[(\mathbf{x} - \boldsymbol{\mu})(\mathbf{y} - \mathbb{E}[\mathbf{y}])^\top\right] = \int \mathbf{x} f(\mathbf{x})^\top \phi(\mathbf{x}; \boldsymbol{\mu}, \Sigma) d\mathbf{x} - \boldsymbol{\mu}\mathbf{m}^\top$$

- The joint distribution of $\mathbf{x}$ and $\mathbf{y}$ can be approximated by a Gaussian:

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu} \\ \mathbf{m} \end{pmatrix}, \begin{bmatrix} \Sigma & C \\ C^\top & S \end{bmatrix} \right)$$

- The approximate distribution of $\mathbf{x}$ conditioned on $\mathbf{y}$ is:

$$\mathbf{x} \mid \mathbf{y} \sim \mathcal{N}\left(\boldsymbol{\mu} + CS^{-1}(\mathbf{y} - \mathbf{m}), \Sigma - CS^{-1}C^\top \right)$$

Nonlinear Kalman Filter Prediction

- Prior state distribution: $\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \Sigma_{t|t})$
- Motion model: $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \qquad \mathbf{w}_t \sim \mathcal{N}(0, W)$
- Force a Gaussian predicted pdf: $\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t})$

$$
\begin{aligned}
\boldsymbol{\mu}_{t+1|t} &= \mathbb{E}\left[\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t}\right] \\
&= \int\int f(\mathbf{x}, \mathbf{u}_t, \mathbf{w}) \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) \phi(\mathbf{w}; 0, W) d\mathbf{x} d\mathbf{w}
\end{aligned}
$$

$$
\begin{aligned}
\Sigma_{t+1|t} &= \mathbb{E}\left[\left(\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t}\right)\left(\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t}\right)^\top \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t}\right] \\
&= \mathbb{E}\left[\mathbf{x}_{t+1}\mathbf{x}_{t+1}^\top \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t}\right] - \boldsymbol{\mu}_{t+1|t}\boldsymbol{\mu}_{t+1|t}^\top \\
&= \int\int f(\mathbf{x}, \mathbf{u}_t, \mathbf{w}) f(\mathbf{x}, \mathbf{u}_t, \mathbf{w})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) \phi(\mathbf{w}; 0, W) d\mathbf{x} d\mathbf{w} - \boldsymbol{\mu}_{t+1|t}\boldsymbol{\mu}_{t+1|t}^\top
\end{aligned}
$$

Nonlinear Kalman Filter Update

- Prior state distribution: $\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t})$
- Observation model: $\mathbf{z}_{t+1} = h(\mathbf{x}_{t+1}, \mathbf{v}_{t+1}), \qquad \mathbf{v}_{t+1} \sim \mathcal{N}(0, V)$
- The Gaussian distribution which approximates the joint distribution of $\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t}$ and $\mathbf{z}_{t+1}$ via moment matching is:

$$\begin{pmatrix} \mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \\ \mathbf{z}_{t+1} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu}_{t+1|t} \\ \mathbf{m}_{t+1|t} \end{pmatrix}, \begin{bmatrix} \Sigma_{t+1|t} & C_{t+1|t} \\ C_{t+1|t}^\top & S_{t+1|t} \end{bmatrix} \right)$$

$$\mathbf{m}_{t+1|t} := \int\int h(\mathbf{x}, \mathbf{v}) \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t}) \phi(\mathbf{v}; 0, V) d\mathbf{x} d\mathbf{v}$$

$$S_{t+1|t} := \int\int (h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})(h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t}) \phi(\mathbf{v}; 0, V) d\mathbf{x} d\mathbf{v}$$

$$C_{t+1|t} := \int\int (\mathbf{x} - \boldsymbol{\mu}_{t+1|t})(h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t}) \phi(\mathbf{v}; 0, V) d\mathbf{x} d\mathbf{v}$$

- The conditional Gaussian distribution of $\mathbf{x}_{t+1} \mid \mathbf{z}_{t+1}, \mathbf{z}_{0:t}, \mathbf{u}_{0:t}$ is then:

$$\begin{aligned} \boldsymbol{\mu}_{t+1|t+1} &:= \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}(\mathbf{z}_{t+1} - \mathbf{m}_{t+1|t}) \\ \Sigma_{t+1|t+1} &:= \Sigma_{t+1|t} - K_{t+1|t} S_{t+1|t} K_{t+1|t}^\top \\ K_{t+1|t} &:= C_{t+1|t} S_{t+1|t}^{-1} \end{aligned}$$

Extended and Unscented Kalman Filters

- The **EKF** and **UKF** use different methods to approximate the five integrals required to implement the nonlinear Kalman filter
- The **EKF** uses a first-order Taylor series approximation to the motion and observation models:

$$f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) \approx f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + \left[\frac{df}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})\right](\mathbf{x}_t - \boldsymbol{\mu}_{t|t}) + \left[\frac{df}{d\mathbf{w}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})\right](\mathbf{w}_t - \mathbf{0})$$

$$h(\mathbf{x}_{t+1}, \mathbf{v}_{t+1}) \approx h(\boldsymbol{\mu}_{t+1|t}, \mathbf{0}) + \left[\frac{dh}{d\mathbf{x}}(\boldsymbol{\mu}_{t+1|t}, \mathbf{0})\right](\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t}) + \left[\frac{dh}{d\mathbf{v}}(\boldsymbol{\mu}_{t+1|t}, \mathbf{0})\right](\mathbf{v}_{t+1} - \mathbf{0})$$

- The **UKF** uses a set of **sigma points** that capture the mean and covariance of the prior Gaussian pdfs to approximate the integrals via a sum. This resembles Monte Carlo approximation but the sigma points are selected **deterministically**.

Extended Kalman Filter Prediction

- Let $F_t := \frac{df}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$ and $Q_t := \frac{df}{d\mathbf{w}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$ so that:

$$f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t) \approx f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t(\mathbf{x}_t - \boldsymbol{\mu}_{t|t}) + Q_t \mathbf{w}_t$$

- Then, the predicted mean and cov can be computed in closed form:

$$\begin{aligned}
\boldsymbol{\mu}_{t+1|t} &\approx \iint \left(f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t(\mathbf{x} - \boldsymbol{\mu}_{t|t}) + Q_t \mathbf{w} \right) \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) \phi(\mathbf{w}; 0, W) d\mathbf{x} d\mathbf{w} \\
&= f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t \left(\int \mathbf{x} \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) d\mathbf{x} - \boldsymbol{\mu}_{t|t} \right) + Q_t \int \mathbf{w} \phi(\mathbf{w}; 0, W) d\mathbf{w} \\
&= \boxed{f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})}
\end{aligned}$$

$$\begin{aligned}
\Sigma_{t+1|t} \approx& \iint \left(f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t(\mathbf{x} - \boldsymbol{\mu}_{t|t}) + Q_t \mathbf{w} \right) \left(f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) + F_t(\mathbf{x} - \boldsymbol{\mu}_{t|t}) + Q_t \mathbf{w} \right)^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) \phi(\mathbf{w}; 0, W) d\mathbf{x} d\mathbf{w} \\
&- \boldsymbol{\mu}_{t+1|t} \boldsymbol{\mu}_{t+1|t}^\top \\
=& f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}) \left(\int (\mathbf{x} - \boldsymbol{\mu}_{t|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) d\mathbf{x} \right) F_t^\top + F_t \left(\int (\mathbf{x} - \boldsymbol{\mu}_{t|t}) \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) d\mathbf{x} \right) f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})^\top \\
&+ F_t \left(\int (\mathbf{x} - \boldsymbol{\mu}_{t|t})(\mathbf{x} - \boldsymbol{\mu}_{t|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t|t}, \Sigma_{t|t}) d\mathbf{x} \right) F_t^\top + Q_t \left(\int \mathbf{w} \mathbf{w}^\top \phi(\mathbf{w}; 0, W) d\mathbf{w} \right) Q_t^\top \\
=& \boxed{F_t \Sigma_{t|t} F_t^\top + Q_t W Q_t^\top}
\end{aligned}$$

Extended Kalman Filter Update

- Let $H_{t+1} := \frac{dh}{d\mathbf{x}}(\boldsymbol{\mu}_{t+1|t}, \mathbf{0})$ and $R_{t+1} := \frac{dh}{d\mathbf{v}}(\boldsymbol{\mu}_{t+1|t}, \mathbf{0})$ so that:

$$h(\mathbf{x}_{t+1}, \mathbf{v}_{t+1}) \approx h(\boldsymbol{\mu}_{t+1|t}, \mathbf{0}) + H_{t+1}(\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t}) + R_{t+1}\mathbf{v}_{t+1}$$

- The joint distribution of $\mathbf{x}_{t+1}$ and $\mathbf{z}_{t+1}$ can be computed in closed form:

$$\mathbf{m}_{t+1|t} := \int\int h(\mathbf{x}, \mathbf{v})\phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t})\phi(\mathbf{v}; 0, V)d\mathbf{x}d\mathbf{v} \approx \boxed{h(\boldsymbol{\mu}_{t+1|t}, \mathbf{0})}$$

$$\begin{aligned} S_{t+1|t} &:= \int\int (h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})(h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t})\phi(\mathbf{v}; 0, V)d\mathbf{x}d\mathbf{v} \\ &\approx \boxed{H_{t+1}\Sigma_{t+1|t}H_{t+1}^\top + R_{t+1}VR_{t+1}^\top} \\ C_{t+1|t} &:= \int\int (\mathbf{x} - \boldsymbol{\mu}_{t+1|t})(h(\mathbf{x}, \mathbf{v}) - \mathbf{m}_{t+1|t})^\top \phi(\mathbf{x}; \boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t})\phi(\mathbf{v}; 0, V)d\mathbf{x}d\mathbf{v} \\ &\approx \boxed{\Sigma_{t+1|t}H_{t+1}^\top} \end{aligned}$$

- The conditional Gaussian distribution of $\mathbf{x}_{t+1} \mid \mathbf{z}_{t+1}$ is then:

$$\begin{aligned} \boldsymbol{\mu}_{t+1|t+1} &:= \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}(\mathbf{z}_{t+1} - \mathbf{m}_{t+1|t}) \\ \Sigma_{t+1|t+1} &:= \Sigma_{t+1|t} - K_{t+1|t}S_{t+1|t}K_{t+1|t}^\top \\ K_{t+1|t} &:= C_{t+1|t}S_{t+1|t}^{-1} \end{aligned}$$

Extended Kalman Filter

Prior: $\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \Sigma_{t|t})$

Motion model:

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \quad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, W)$$

$$F_t := \frac{df}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0}), \quad Q_t := \frac{df}{d\mathbf{w}}(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$$

Obs. model:

$$\mathbf{z}_t = h(\mathbf{x}_t, \mathbf{v}_t), \quad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, V)$$

$$H_t := \frac{dh}{d\mathbf{x}}(\boldsymbol{\mu}_{t|t-1}, \mathbf{0}), \quad R_t := \frac{dh}{d\mathbf{v}}(\boldsymbol{\mu}_{t|t-1}, \mathbf{0})$$

Prediction:

$$\boldsymbol{\mu}_{t+1|t} = f(\boldsymbol{\mu}_{t|t}, \mathbf{u}_t, \mathbf{0})$$

$$\Sigma_{t+1|t} = F_t \Sigma_{t|t} F_t^\top + Q_t W Q_t^\top$$

Update:

$$\mu_{t+1|t+1} = \mu_{t+1|t} + K_{t+1|t}(z_{t+1} - h(\mu_{t+1|t}, 0))$$

$$\Sigma_{t+1|t+1} = (I - K_{t+1|t} H_{t+1})\Sigma_{t+1|t}$$

Kalman Gain: $K_{t+1|t} := \Sigma_{t+1|t} H_{t+1}^\top \left(H_{t+1} \Sigma_{t+1|t} H_{t+1}^\top + R_{t+1} V R_{t+1}^\top \right)^{-1}$

Unscented Transform

- The **unscented transform** (Julier et al., 1995, 2000) is a numerical method for approximating the joint distribution of a Gaussian random variable $\mathbf{x} \in \mathbb{R}^d$ and a nonlinear transformation f of it:

$$\mathbf{y} = f(\mathbf{x}), \qquad \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu} \\ \mathbf{m} \end{pmatrix}, \begin{bmatrix} \Sigma & C \\ C^\top & S \end{bmatrix} \right)$$

- Choose a set of $2d+1$ **sigma points** using the i-th columns of the square root $\sqrt{\Sigma}$ of the covariance $\Sigma = \sqrt{\Sigma}\sqrt{\Sigma}^\top$ (**note**: $\sqrt{\Sigma}$ is lower-triangular and can be obtained, e.g., via **Cholesky factorization**):

$$\mathcal{X}^{(0)} = \boldsymbol{\mu}, \qquad \mathcal{X}^{(i)} = \boldsymbol{\mu} \pm \sqrt{d+\lambda}\left[\sqrt{\Sigma}\right]_i, \quad i = 1, \dots, d$$

$$\lambda := \alpha^2(d+k) - d, \qquad (\alpha, k) : \text{determine sigma points spread, e.g.,}$$

$$\alpha \in [0.0001, 1], k = 0$$

Unscented Transform

- The sigma points capture the shape of the original distribution of $\mathbf{x}$ well and can be propagated through the nonlinear function f to estimate the mean and covariance of the transformed variable

- The mean and covariance of $\mathbf{y} = f(\mathbf{x})$ are estimated using the sigma points $\mathcal{X}^{(0)} = \boldsymbol{\mu}$ and $\mathcal{X}^{(i)} = \boldsymbol{\mu} \pm \sqrt{d+\lambda}\left[\sqrt{\Sigma}\right]_i$, $i = 1, \ldots, d$:

$$\mathbf{m} = \sum_{i=0}^{2d} W_i^{(m)} f\left(\mathcal{X}^{(i)}\right), \qquad W_0^{(m)} = \frac{\lambda}{d+\lambda}, \qquad W_i^{(m)} = \frac{1}{2(d+\lambda)}, \quad i = 1, \ldots, 2d$$

$$S = \sum_{i=0}^{2d} W_i^{(c)} \left(f\left(\mathcal{X}^{(i)}\right) - \mathbf{m}\right)\left(f\left(\mathcal{Y}^{(i)}\right) - \mathbf{m}\right)^\top, \qquad W_0^{(c)} = \frac{\lambda}{d+\lambda} + (1 - \alpha^2 + \beta)$$

e.g., $\alpha \in [0.0001, 1]$, $\beta \in \{0, 2\}$

$$C = \sum_{i=0}^{2d} W_i^{(c)} \left(\mathcal{X}^{(i)} - \boldsymbol{\mu}\right)\left(f\left(\mathcal{X}^{(i)}\right) - \mathbf{m}\right)^\top, \qquad W_i^{(c)} = \frac{1}{2(d+\lambda)}, \quad i = 1, \ldots, 2d$$

Unscented Kalman Filter Prediction

- **Prior**: $\mathbf{x}_t \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t-1} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \Sigma_{t|t})$
- **Motion Model**: $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \mathbf{w}_t), \quad \mathbf{w}_t \sim \mathcal{N}(0, W)$

$$\begin{pmatrix} \mathcal{X}^{(0)}_{t|t} \\ \mathcal{W}^{(0)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{t|t} \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} \mathcal{X}^{(i)}_{t|t} \\ \mathcal{W}^{(i)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{t|t} \\ 0 \end{pmatrix} \pm \sqrt{(d_x + d_w)} \begin{bmatrix} \sqrt{\Sigma_{t|t}} & 0 \\ 0 & \sqrt{W} \end{bmatrix}_i$$

- **Prediction**: for $\alpha = 1$, $k = 0$, $\lambda = 0$, and $\beta = 2$

$$\boldsymbol{\mu}_{t+1|t} = \sum_{i=0}^{2(d_x+d_w)} W_i^{(m)} f\left(\mathcal{X}^{(i)}_{t|t}, \mathbf{u}_t, \mathcal{W}^{(i)}\right)$$

$$\Sigma_{t+1|t} = \sum_{i=0}^{2(d_x+d_w)} W_i^{(c)} \left(f\left(\mathcal{X}^{(i)}_{t|t}, \mathbf{u}_t, \mathcal{W}^{(i)}\right) - \boldsymbol{\mu}_{t+1|t}\right)\left(f\left(\mathcal{X}^{(i)}_{t|t}, \mathbf{u}_t, \mathcal{W}^{(i)}\right) - \boldsymbol{\mu}_{t+1|t}\right)^\top$$

- **Weights**:

$$W_0^{(m)} = 0, \qquad W_i^{(m)} = \frac{1}{2(d_x + d_w)}, \quad i = 1, \ldots, 2(d_x + d_w)$$

$$W_0^{(c)} = 2, \qquad W_i^{(c)} = \frac{1}{2(d_x + d_w)}, \quad i = 1, \ldots, 2(d_x + d_w)$$

Unscented Kalman Filter Update

- **Prior**: $\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t})$
- **Observation Model**: $\mathbf{z}_{t+1} = h(\mathbf{x}_{t+1}, \mathbf{v}_{t+1}), \quad \mathbf{v}_{t+1} \sim \mathcal{N}(0, V)$

$$\begin{pmatrix} \mathcal{X}^{(0)}_{t+1|t} \\ \mathcal{V}^{(0)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{t+1|t} \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} \mathcal{X}^{(i)}_{t+1|t} \\ \mathcal{V}^{(i)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_{t+1|t} \\ 0 \end{pmatrix} \pm \sqrt{(d_x + d_v)} \begin{bmatrix} \sqrt{\Sigma_{t+1|t}} & 0 \\ 0 & \sqrt{V} \end{bmatrix}_i$$

- **Update**: $$\begin{aligned} \boldsymbol{\mu}_{t+1|t+1} &= \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}(\mathbf{z}_{t+1} - \mathbf{m}_{t+1|t})) \\ \Sigma_{t+1|t+1} &= \Sigma_{t+1|t} - K_{t+1|t} S_{t+1|t} K^\top_{t+1|t} \end{aligned}$$
- **Kalman Gain**: $K_{t+1|t} = C_{t+1|t} S^{-1}_{t+1|t}$

$$\mathbf{m}_{t+1|t} = \sum_{i=0}^{2(d_x+d_v)} W_i^{(m)} h\left(\mathcal{X}^{(i)}_{t+1|t}, \mathcal{V}^{(i)}\right)$$

$$S_{t+1|t} = \sum_{i=0}^{2(d_x+d_v)} W_i^{(c)} \left(h\left(\mathcal{X}^{(i)}_{t+1|t}, \mathcal{V}^{(i)}\right) - \mathbf{m}_{t+1|t}\right)\left(h\left(\mathcal{X}^{(i)}_{t+1|t}, \mathcal{V}^{(i)}\right) - \mathbf{m}_{t+1|t}\right)^\top$$

$$C_{t+1|t} = \sum_{i=0}^{2(d_x+d_v)} W_i^{(c)} \left(\mathcal{X}^{(i)}_{t+1|t} - \boldsymbol{\mu}_{t+1|t}\right)\left(h\left(\mathcal{X}^{(i)}_{t+1|t}, \mathcal{V}^{(i)}\right) - \mathbf{m}_{t+1|t}\right)^\top$$

Unscented Kalman Filter (additive noise)

Prior $\mathbf{x}_{t+1} \mid \mathbf{z}_{0:t}, \mathbf{u}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t+1|t}, \Sigma_{t+1|t})$

Motion model $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t) + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(0, W)$

Obs. model $\mathbf{z}_{t+1} = h(\mathbf{x}_{t+1}) + \mathbf{v}_{t+1}, \quad \mathbf{v}_{t+1} \sim \mathcal{N}(0, V)$

Predict

$$\boldsymbol{\mu}_{t+1|t} = \sum_{i=0}^{2d_x} W_i^{(m)} f\left(\mathcal{X}_{t|t}^{(i)}, \mathbf{u}_t\right), \quad \mathcal{X}_{t|t}^{(0)} = \boldsymbol{\mu}_{t|t}, \quad \mathcal{X}_{t|t}^{(i)} = \boldsymbol{\mu}_{t|t} \pm \sqrt{d_x + \lambda}\left[\sqrt{\Sigma_{t|t}}\right]_i$$

$$\Sigma_{t+1|t} = \sum_{i=0}^{2d_x} W_i^{(c)} \left(f\left(\mathcal{X}_{t|t}^{(i)}, \mathbf{u}_t\right) - \boldsymbol{\mu}_{t+1|t}\right)\left(f\left(\mathcal{X}_{t|t}^{(i)}, \mathbf{u}_t\right) - \boldsymbol{\mu}_{t+1|t}\right)^\top + W$$

Update

$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}(\mathbf{z}_{t+1} - \mathbf{m}_{t+1|t})$$

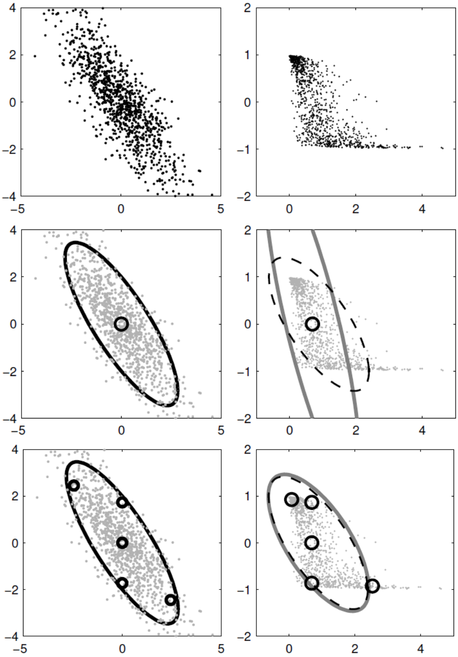
$$\Sigma_{t+1|t+1} = \Sigma_{t+1|t} - K_{t+1|t} S_{t+1|t} K_{t+1|t}^\top$$

Kalman gain $K_{t+1|t} = C_{t+1|t} S_{t+1|t}^{-1}$

$$\mathbf{m}_{t+1|t} = \sum_{i=0}^{2d_x} W_i^{(m)} h\left(\mathcal{X}_{t+1|t}^{(i)}\right), \qquad \mathcal{X}_{t+1|t}^{(0)} = \boldsymbol{\mu}_{t+1|t}, \quad \mathcal{X}_{t+1|t}^{(i)} = \boldsymbol{\mu}_{t+1|t} \pm \sqrt{d_x + \lambda}\left[\sqrt{\Sigma_{t+1|t}}\right]_i$$

$$S_{t+1|t} = \sum_{i=0}^{2d_x} W_i^{(c)} \left(h\left(\mathcal{X}_{t+1|t}^{(i)}\right) - \mathbf{m}_{t+1|t}\right)\left(h\left(\mathcal{X}_{t+1|t}^{(i)}\right) - \mathbf{m}_{t+1|t}\right)^\top + V$$

$$C_{t+1|t} = \sum_{i=0}^{2d_x} W_i^{(c)} \left(\mathcal{X}_{t+1|t}^{(i)} - \boldsymbol{\mu}_{t+1|t}\right)\left(h\left(\mathcal{X}_{t+1|t}^{(i)}\right) - \mathbf{m}_{t+1|t}\right)^\top$$



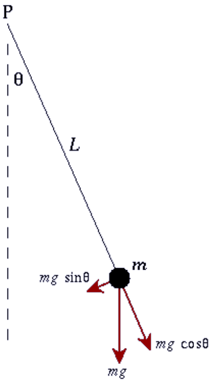
Applying a nonlinear transformation to the random variable on the left results in a new random variable (right)

EKF: Linearization-based approximation to the transformation formed by calculating the curvature at the mean (solid) and true covariance (dashed).

UKF: Unscented transform approximation to the transformation formed by propagating sigma points through the nonlinearity and estimating the covariance from them (solid) and true covariance (dashed).

Noisy Pendulum Tracking

- Consider a simple pendulum consisting of a mass m hanging from a string of length L and fixed at a pivot point P



- The differential equation for the pendulum motion can be obtained using Newton's second law for rotational systems which relates the net external torque τ (position $\times$ force) to the product of the moment of inertia $I = mL^2$ and the angular acceleration $\ddot{\theta}(t)$:

$$\tau = -mgL\sin\theta(t) = mL^2\ddot{\theta}(t) \quad \Rightarrow \quad \ddot{\theta}(t) = -\frac{g}{L}\sin\theta(t) + \underbrace{w(t)}_{\text{noise}\sim\mathcal{N}(0,q)}$$

- The model can be converted into a state-space model with state $\mathbf{x}(t) := (\theta(t), \omega(t))^\top$, where $\omega(t) := \dot{\theta}(t)$ as follows:

$$\frac{d}{dt}\begin{pmatrix}\theta(t)\\ \omega(t)\end{pmatrix} = \begin{bmatrix}\omega(t)\\ -\frac{g}{L}\sin(\theta(t))\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix} w(t)$$

Discrete-time Model

- **Motion model**: a simple discretization of the pendulum state-space model with sampling period τ leads to:

$$\mathbf{x}_{t+1} = \begin{pmatrix} \theta_{t+1} \\ \omega_{t+1} \end{pmatrix} = \underbrace{\begin{bmatrix} \theta_t + \tau\omega_t \\ \omega_t - \tau\frac{g}{L}\sin\theta_t \end{bmatrix}}_{f(\mathbf{x}_t)} + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}\left(0, \underbrace{q\begin{bmatrix} \frac{\tau^3}{3} & \frac{\tau^2}{2} \\ \frac{\tau^2}{2} & \tau \end{bmatrix}}_{W}\right)$$

- **Observation model**: consider estimating the angle θ_t and the velocity ω_t of the pendulum using measurements of its deviation from rest position, i.e.,:

$$z_t = \underbrace{L\sin(\theta_t)}_{h(x_t)} + v_t, \qquad v_t \sim \mathcal{N}(0, V)$$

Extended Kalman Filter

- **Jacobians** of the motion and observation model for $\mathbf{x} = (\theta, \omega)^\top$:

$$F(\mathbf{x}) := \begin{bmatrix} 1 & \tau \\ -\tau \frac{g}{L} \cos\theta & 1 \end{bmatrix} \qquad H(\mathbf{x}) = \begin{bmatrix} L\cos\theta & 0 \end{bmatrix}$$

- **Prior**: $\mathbf{x}_t \mid \mathbf{z}_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \Sigma_{t|t})$

- **Prediction**:

$$\boldsymbol{\mu}_{t+1|t} = f(\boldsymbol{\mu}_{t|t}) = \begin{bmatrix} \mu^{\theta}_{t|t} + \tau \mu^{\omega}_{t|t} \\ \mu^{\omega}_{t|t} - \tau \frac{g}{L} \sin \mu^{\theta}_{t|t} \end{bmatrix}$$

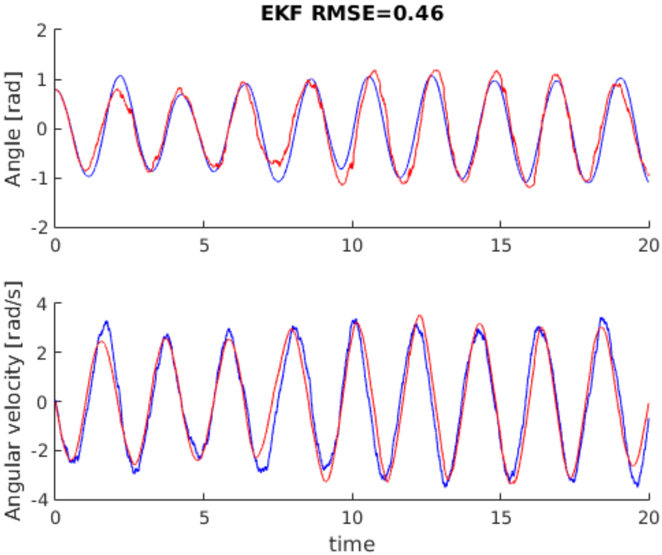
$$\Sigma_{t+1|t} = F(\mu_{t|t}) \Sigma_{t|t} F(\mu_{t|t})^\top + W$$

Extended Kalman Filter

- **Innovation**: $\nu_{t+1|t} := z_{t+1} - L\sin(\mu^{\theta}_{t+1|t})$
- **Measurement/innovation covariance**:
  $S_{t+1|t} := H(\mu_{t+1|t})\Sigma_{t+1|t}H(\mu_{t+1|t})^{\top} + V$
- **State-measurement cross-covariance**: $\Sigma_{t+1|t}H(\mu_{t+1|t})^{\top}$
- **Kalman gain**: $K_{t+1|t} = \Sigma_{t+1|t}H(\mu_{t+1|t})^{\top}S^{-1}_{t+1|t}$
- **Update**: $\begin{aligned}\boldsymbol{\mu}_{t+1|t+1} &= \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}\nu_{t+1|t}\\ \Sigma_{t+1|t+1} &= \Sigma_{t+1|t} - K_{t+1|t}H(\mu_{t+1|t})\Sigma_{t+1|t}\end{aligned}$

EKF Performance

- $\tau = 0.001$, $q = 0.3$, $g = 9.81$, $L = 1$, $V = 0.64$
- Prediction at 1000 Hz, update at 20 Hz



Uscented Kalman Filter Prediction

- **Prior**: $\mathbf{x}_t \mid z_{0:t} \sim \mathcal{N}(\boldsymbol{\mu}_{t|t}, \Sigma_{t|t})$
- **Sigma points** with parameter $\lambda = \alpha^2(d_x + k) - d_x$ determining the sigma point spread (usual choice: $\alpha \in [0.0001, 1]$, $k = 0$, $\beta \in \{0, 2\}$)

$$\mathcal{X}_{t|t}^{(0)} = \boldsymbol{\mu}_{t|t}, \qquad \mathcal{X}_{t|t}^{(i)} = \boldsymbol{\mu}_{t|t} \pm \sqrt{(d_x + \lambda)} \left[\sqrt{\Sigma_{t|t}}\right]_i, \quad i = 1, \ldots, 2d_x$$

$$W_m^{(0)} = \frac{\lambda}{d_x + \lambda}, \qquad W_m^{(i)} = \frac{1}{2(d_x + \lambda)}, \quad i = 1, \ldots, 2d_x$$

$$W_c^{(0)} = \frac{\lambda}{d_x + \lambda} + (1 - \alpha^2 + \beta), \qquad W_c^{(i)} = \frac{1}{2(d_x + \lambda)}, \quad i = 1, \ldots, 2d_x$$

- **Prediction**:

$$\boldsymbol{\mu}_{t+1|t} = \sum_{i=0}^{2d_x} W_m^{(i)} f\left(\mathcal{X}_{t|t}^{(i)}\right)$$

$$\Sigma_{t+1|t} = \sum_{i=0}^{2d_x} W_c^{(i)} \left(f\left(\mathcal{X}_{t|t}^{(i)}\right) - \boldsymbol{\mu}_{t+1|t}\right)\left(f\left(\mathcal{X}_{t|t}^{(i)}\right) - \boldsymbol{\mu}_{t+1|t}\right)^\top + W$$

Uscented Kalman Filter Update

- **Sigma points**:

$$\mathcal{X}^{(0)}_{t+1|t} = \boldsymbol{\mu}_{t+1|t}, \qquad \mathcal{X}^{(i)}_{t+1|t} = \boldsymbol{\mu}_{t+1|t} \pm \sqrt{(d_x+\lambda)}\left[\sqrt{\Sigma_{t+1|t}}\right]_i, \quad i = 1,\ldots,2d_x$$

- **Expected measurement**: $m_{t+1|t} = \sum_{i=0}^{2d_x} W_m^{(i)} h\left(\mathcal{X}^{(i)}_{t+1|t}\right)$
- **Innovation**: $\nu_{t+1|t} := z_{t+1} - m_{t+1|t}$
- **Measurement/innovation covariance**:

$$S_{t+1|t} = \sum_{i=0}^{2d_x} W_c^{(i)} \left(h\left(\mathcal{X}^{(i)}_{t+1|t}\right) - m_{t+1|t}\right)\left(h\left(\mathcal{X}^{(i)}_{t+1|t}\right) - m_{t+1|t}\right)^\top + V$$

- **State-measurement cross-covariance**:

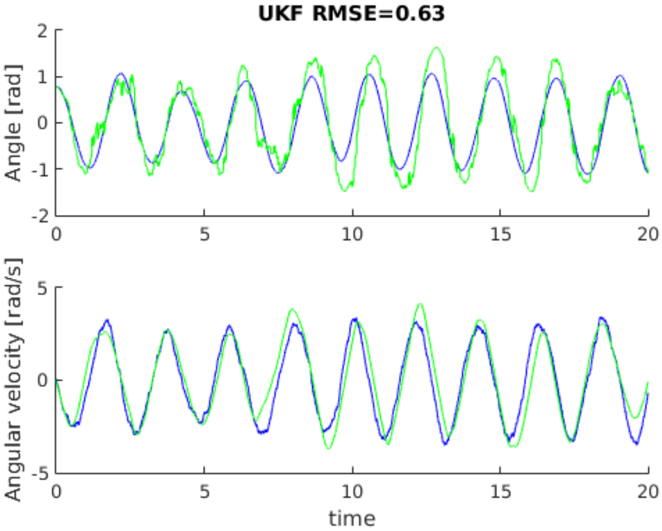
$$C_{t+1|t} = \sum_{i=0}^{2d_x} W_c^{(i)} \left(\mathcal{X}^{(i)}_{t+1|t} - \boldsymbol{\mu}_{t+1|t}\right)\left(h\left(\mathcal{X}^{(i)}_{t+1|t}\right) - m_{t+1|t}\right)^\top$$

- **Kalman gain**: $K_{t+1|t} = C_{t+1|t} S^{-1}_{t+1|t}$
- **Update**:

$$\begin{aligned}\boldsymbol{\mu}_{t+1|t+1} &= \boldsymbol{\mu}_{t+1|t} + K_{t+1|t}\nu_{t+1|t}\\ \Sigma_{t+1|t+1} &= \Sigma_{t+1|t} - K_{t+1|t} S_{t+1|t} K^\top_{t+1|t}\end{aligned}$$

UKF Performance

- $\tau = 0.001$, $q = 0.3$, $g = 9.81$, $L = 1$, $V = 0.64$
- Prediction at 1000 Hz, update at 20 Hz



UKF vs EKF Predicted Covariance

- Prior: $\mathcal{N}\left(\begin{pmatrix}\frac{\pi}{4} \\ -1\end{pmatrix}, \begin{bmatrix}2 & -0.3 \\ -0.3 & 0.5\end{bmatrix}\right)$
- One prediction step with parameters $\tau = 1$, $g = 9.81$, $L = 1$