Module 2-3: Introduction to Stochastic Control

Linear Control Systems (2020)

Ding Zhao

Assistant Professor

College of Engineering

School of Computer Science

Carnegie Mellon University

Table of Contents

Table of Contents

Plan for Today

- Prediction - Correction Observer (Luenberger Observer)
- Prediction - Updated-Correction Observer
- Kalman Filter

All observers have the same goal:

$$y(k), u(k), \hat{x}(k-1) \Rightarrow \hat{x}(k)$$

Recap: Continuous Time Luenberger Observer

By using an observer and performing state feedback on $\hat{x}$, we can build an output feedback controller.

$$\dot{\bar{x}} = A\hat{x} + Bu \triangleleft \text{ prediction}$$

$$\dot{\hat{x}} = \dot{\bar{x}} + L(y - \hat{y}) \triangleleft \text{ correction}$$

$$u = K\hat{x} \qquad\qquad y = Cx + Du$$

$$\Rightarrow \dot{x} = Ax + BK\hat{x} \qquad\qquad \hat{y} = C\hat{x} + Du$$

$$\begin{aligned} \Rightarrow \dot{\hat{x}} &= A\hat{x} + BK\hat{x} + L(Cx + Du - (C\hat{x} + Du)) \\ &= (A - LC + BK)\hat{x} + LCx \end{aligned}$$

- Now stack $\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} A & BK \\ LC & A - LC + BK \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} \Rightarrow \begin{bmatrix} \dot{x} \\ \dot{e} \end{bmatrix} = \begin{bmatrix} A + BK & -BK \\ 0 & A - LC \end{bmatrix} \begin{bmatrix} x \\ e \end{bmatrix}$
- Note D does not show up in the final results. Some analysis actually takes D as 0. It makes sense because we can also take $\tilde{y} = y - Du = Cx$.

Recap: Discrete Luenberger Observer (in Pole Placement Notation)

Controller	System
$\Rightarrow \hat{x}(k-1)$	$\leftarrow y(k-1)$
$u(k) = K\hat{x}(k-1)$	$\rightarrow u(k)$
$\bar{x}(k) = A\hat{x}(k-1) + Bu(k)$	$x(k) = Ax(k-1) + Bu(k)$
$\hat{y}(k-1) = C\hat{x}(k-1)$	
$\hat{x}(k) = \bar{x}(k) + L(y(k-1) - \hat{y}(k-1))$	
wait until ΔT	
$\Rightarrow \hat{x}(k)$	$\leftarrow y(k) = Cx(k)$
Memorize $\hat{x}(k)$	
Output $u(k+1) = K\hat{x}(k)$	$\rightarrow u(k+1)$
wait until ΔT	

Note we do not use $\hat{x}(k)$ until next time interval.

A Similar Expression

Controller	System
$\Rightarrow \hat{x}(k-1)$	$\leftarrow y(k-1)$
$u(k) = K\hat{x}(k-1)$	$\rightarrow u(k)$
	$x(k) = Ax(k-1) + Bu(k)$

wait until ΔT	
$\Rightarrow \hat{x}(k-1)$	$\leftarrow y(k) = Cx(k)$
$\bar{x}(k) = A\hat{x}(k-1) + Bu(k)$	$\triangleleft$ We could compute the red part
$\hat{y}(k-1) = C\hat{x}(k-1)$	in the next time interval
$\hat{x}(k) = \bar{x}(k) + L(y(k-1) - \hat{y}(k-1))$	
Memorize $\hat{x}(k)$	
Output $u(k+1) = K\hat{x}(k)$	$\rightarrow u(k+1)$

wait until ΔT

Prediction Updated-Correction (in Pole Placement Notation)

Controller	System
$\Rightarrow \hat{x}(k-1)$	$\leftarrow y(k-1)$
$u(k) = K\hat{x}(k-1)$	$\rightarrow u(k)$
	$x(k) = Ax(k-1) + Bu(k)$
wait until ΔT	
$\Rightarrow \hat{x}(k-1)$	$\leftarrow y(k) = Cx(k)$
$\bar{x}(k) = A\hat{x}(k-1) + Bu(k)$	
$\hat{y}(k) = C\bar{x}(k)$	
$\hat{x}(k) = \bar{x}(k) + L(y(k) - \hat{y}(k))$	$\triangleleft$ Since $y(k)$ has been available, use it instead.
Memorize $\hat{x}(k)$	
Output $u(k+1) = K\hat{x}(k)$	$\rightarrow u(k+1)$
wait until ΔT	

Prediction Updated-Correction (in Kalman Filter Notation)

Controller	System
$\Rightarrow \hat{\mathbf{x}}_{k-1\|k-1}$	$\leftarrow \mathbf{y}_{k-1}$
$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_{k-1\|k-1}$	$\rightarrow \mathbf{u}_k$
	$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k$
wait until ΔT	
$\Rightarrow \hat{\mathbf{x}}_{k-1\|k-1}$	$\leftarrow \mathbf{y}_k = \mathbf{C}\mathbf{x}_k$
$\hat{\mathbf{x}}_{k\|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1\|k-1} + \mathbf{B}\mathbf{u}_k$	
$\hat{\mathbf{y}}_{k\|k-1} = \mathbf{C}\hat{\mathbf{x}}_{k\|k-1}$	
$\hat{\mathbf{x}}_{k\|k} = \hat{\mathbf{x}}_{k\|k-1} + \mathbf{L}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k\|k-1})$	
Memorize $\hat{\mathbf{x}}_{k\|k}$	
Output $\mathbf{u}_{k+1} = \mathbf{K}\hat{\mathbf{x}}_{k\|k}$	$\rightarrow \mathbf{u}_{k+1}$
wait until ΔT	

Prediction Updated-Correction (Consider Gaussian Noise)

Controller	System
$\Rightarrow \hat{\mathbf{x}}_{k-1\|k-1}$	$\leftarrow \mathbf{y}_{k-1}$
$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_{k-1\|k-1}$	$\rightarrow \mathbf{u}_k$
	$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k$
wait until ΔT	
$\Rightarrow \hat{\mathbf{x}}_{k-1\|k-1}$	$\leftarrow \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k$
$\hat{\mathbf{x}}_{k\|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1\|k-1} + \mathbf{B}\mathbf{u}_k$	
$\hat{\mathbf{y}}_{k\|k-1} = \mathbf{C}\hat{\mathbf{x}}_{k\|k-1}$	
$\hat{\mathbf{x}}_{k\|k} = \hat{\mathbf{x}}_{k\|k-1} + \mathbf{L}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k\|k-1})$	
Memorize $\hat{\mathbf{x}}_{k\|k}$	
Output $\mathbf{u}_{k+1} = \mathbf{K}\hat{\mathbf{x}}_{k\|k}$	$\rightarrow \mathbf{u}_{k+1}$
wait until ΔT	

$\mathbf{w}_k$ is the process noise, $\mathbf{v}_k$ is the observation noise. $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{W})$, $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{V})$.

Prediction Updated-Correction (with Kalman Optimal Gain)

Controller	System
$\Rightarrow \hat{\mathbf{x}}_{k-1\|k-1}$	$\leftarrow \mathbf{y}_{k-1}$
$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_{k-1\|k-1}$	$\rightarrow \mathbf{u}_k$
	$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k$
wait until ΔT	
$\Rightarrow \hat{\mathbf{x}}_{k-1\|k-1}, \mathbf{P}_{k-1\|k-1}$	$\leftarrow \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k$
$\hat{\mathbf{x}}_{k\|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1\|k-1} + \mathbf{B}\mathbf{u}_k$	
$\hat{\mathbf{y}}_{k\|k-1} = \mathbf{C}\hat{\mathbf{x}}_{k\|k-1}$	
$\mathbf{P}_{k\|k-1} = \mathbf{A}\mathbf{P}_{k-1\|k-1}\mathbf{A}^{\mathrm{T}} + \mathbf{W}$	
$\mathbf{L}_k = \mathbf{P}_{k\|k-1}\mathbf{C}^{\mathrm{T}}(\mathbf{C}\mathbf{P}_{k\|k-1}\mathbf{C}^{\mathrm{T}} + \mathbf{V})^{-1}$	
$\hat{\mathbf{x}}_{k\|k} = \hat{\mathbf{x}}_{k\|k-1} + \mathbf{L}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k\|k-1})$	
$\mathbf{P}_{k\|k} = (\mathbf{I} - \mathbf{L}_k\mathbf{C})\mathbf{P}_{k\|k-1}$	
Memorize $\hat{\mathbf{x}}_{k\|k}, \mathbf{P}_{k\|k}$	
Output $\mathbf{u}_{k+1} = \mathbf{K}\hat{\mathbf{x}}_{k\|k}$	$\rightarrow \mathbf{u}_{k+1}$
wait until ΔT	

The Five Kalman Filter Equations

Controller	System
$\Rightarrow \hat{\mathbf{x}}_{k-1\|k-1}$	$\leftarrow \mathbf{y}_{k-1}$
$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_{k-1\|k-1}$	$\rightarrow \mathbf{u}_k$
	$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k$
wait until ΔT	
$\Rightarrow \hat{\mathbf{x}}_{k-1\|k-1}, \mathbf{P}_{k-1\|k-1}$	$\leftarrow \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k$
P-1: $\hat{\mathbf{x}}_{k\|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1\|k-1} + \mathbf{B}\mathbf{u}_k$	
$\hat{\mathbf{y}}_{k\|k-1} = \mathbf{C}\hat{\mathbf{x}}_{k\|k-1}$	
P-2: $\mathbf{P}_{k\|k-1} = \mathbf{A}\mathbf{P}_{k-1\|k-1}\mathbf{A}^{\mathrm{T}} + \mathbf{W}$	
C-1: $\mathbf{L}_k = \mathbf{P}_{k\|k-1}\mathbf{C}^{\mathrm{T}}(\mathbf{C}\mathbf{P}_{k\|k-1}\mathbf{C}^{\mathrm{T}} + \mathbf{V})^{-1}$	
C-2: $\hat{\mathbf{x}}_{k\|k} = \hat{\mathbf{x}}_{k\|k-1} + \mathbf{L}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k\|k-1})$	
C-3: $\mathbf{P}_{k\|k} = (\mathbf{I} - \mathbf{L}_k\mathbf{C})\mathbf{P}_{k\|k-1}$	
Memorize $\hat{\mathbf{x}}_{k\|k}, \mathbf{P}_{k\|k}$	
Output $\mathbf{u}_{k+1} = \mathbf{K}\hat{\mathbf{x}}_{k\|k}$	$\rightarrow \mathbf{u}_{k+1}$
wait until ΔT	

The Five Kalman Filter Equations II

Controller	System
$\Rightarrow \hat{\mathbf{x}}_{k-1\|k-1}$	$\leftarrow \mathbf{y}_{k-1}$
$\mathbf{u}_k = \mathbf{K}\hat{\mathbf{x}}_{k-1\|k-1}$	$\rightarrow \mathbf{u}_k$
	$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k$
wait until ΔT	
$\Rightarrow \hat{\mathbf{x}}_{k-1\|k-1}, \mathbf{P}_{k-1\|k-1}$	$\leftarrow \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k$
P-1: $\hat{\mathbf{x}}_{k\|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1\|k-1} + \mathbf{B}\mathbf{u}_k$	
P-2: $\mathbf{P}_{k\|k-1} = \mathbf{A}\mathbf{P}_{k-1\|k-1}\mathbf{A}^{\mathrm{T}} + \mathbf{W}$	
C-1: $\mathbf{L}_k = \mathbf{P}_{k\|k-1}\mathbf{C}^{\mathrm{T}}(\mathbf{C}\mathbf{P}_{k\|k-1}\mathbf{C}^{\mathrm{T}} + \mathbf{V})^{-1}$	
C-2: $\hat{\mathbf{x}}_{k\|k} = \hat{\mathbf{x}}_{k\|k-1} + \mathbf{L}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_{k\|k-1})$	
C-3: $\mathbf{P}_{k\|k} = (\mathbf{I} - \mathbf{L}_k\mathbf{C})\mathbf{P}_{k\|k-1}$	
Memorize $\hat{\mathbf{x}}_{k\|k}, \mathbf{P}_{k\|k}$	
Output $\mathbf{u}_{k+1} = \mathbf{K}\hat{\mathbf{x}}_{k\|k}$	$\rightarrow \mathbf{u}_{k+1}$
wait until ΔT	

How to Compute the Kalman Gain $\mathbf{L}_k$

$$\min_{\mathbf{L}_k} \mathrm{E}\left[\left\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right\|^2\right]$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})$$

s.t. $\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}\mathbf{u}_k$

$\hat{\mathbf{y}}_{k|k-1} = \mathbf{C}\hat{\mathbf{x}}_{k|k-1}$

The system dynamics:

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k$$
$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k$$

$\mathbf{w}_k$ is the process noise, $\mathbf{v}_k$ is the observation noise. $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{W})$, $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{V})$.

Kalman Filter:

P-1: $\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}\mathbf{u}_k$

P-2: $\mathbf{P}_{k|k-1} = \mathbf{A}\mathbf{P}_{k-1|k-1}\mathbf{A}^{\mathrm{T}} + \mathbf{W}$

C-1: $\mathbf{L}_k = \mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}}(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}} + \mathbf{V})^{-1}$

C-2: $\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1})$

C-3: $\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{L}_k\mathbf{C})\mathbf{P}_{k|k-1}$

Where do these three equations come from?

The Key Ideas of Kalman Filter

- What are $\mathbf{P}_{k|k}$ and $\mathbf{P}_{k|k-1}$

$$\begin{aligned}
\mathbf{x}_k &= \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \\
\hat{\mathbf{x}}_{k|k-1} &= \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}\mathbf{u}_k \\
\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1}) \\
\mathbf{P}_{k|k-1} &= \mathrm{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \\
\mathbf{P}_{k|k} &= \mathrm{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})
\end{aligned}$$

They are Covariance Matrices

Covariance Matrix

Let random vector $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$. The covariance matrix of $\mathbf{x}$ is defined as

$$\mathrm{cov}(\mathbf{x}) = \mathrm{E}[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T] = \mathrm{E}(\mathbf{x}\mathbf{x}^\mathrm{T}) - \bar{\mathbf{x}}\bar{\mathbf{x}}^\mathrm{T}$$

$$= \begin{bmatrix} \mathrm{E}[(x_1 - \bar{x}_1)(x_1 - \bar{x}_1)] & \mathrm{E}[(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)] & \cdots & \mathrm{E}[(x_1 - \bar{x}_1)(x_n - \bar{x}_n)] \\ \mathrm{E}[(x_2 - \bar{x}_2)(x_1 - \bar{x}_1)] & \mathrm{E}[(x_2 - \bar{x}_2)(x_2 - \bar{x}_2)] & \cdots & \mathrm{E}[(x_2 - \bar{x}_2)(x_n - \bar{x}_n)] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}[(x_n - \bar{x}_n)(x_1 - \bar{x}_1)] & \mathrm{E}[(x_n - \bar{x}_n)(x_2 - \bar{x}_2)] & \cdots & \mathrm{E}[(x_n - \bar{x}_n)(x_n - \bar{x}_n)] \end{bmatrix}$$

$\mathrm{cov}(\mathbf{A}\mathbf{x} + \mathbf{a}) = \mathbf{A}\ \mathrm{cov}(\mathbf{x})\ \mathbf{A}^\mathrm{T}$, where $\mathbf{A}$ and $\mathbf{a}$ are constant.
$\mathrm{cov}(\mathbf{x} + \mathbf{y}) = \mathrm{cov}(\mathbf{x}) + \mathrm{cov}(\mathbf{x}, \mathbf{y}) + \mathrm{cov}(\mathbf{y}, \mathbf{x}) + \mathrm{cov}(\mathbf{y})$, where $\mathrm{cov}(\mathbf{x}, \mathbf{y})$ is the cross-covariance matrix of $\mathbf{x}$ and $\mathbf{y}$. If $\mathbf{x}$ and $\mathbf{y}$ are independent, $\mathrm{cov}(\mathbf{x}, \mathbf{y}) = \mathrm{cov}(\mathbf{y}, \mathbf{x}) = \mathbf{0}$
$\Rightarrow \mathrm{cov}(\mathbf{x} + \mathbf{y}) = \mathrm{cov}(\mathbf{x}) + \mathrm{cov}(\mathbf{y})$

The Key Ideas of Kalman Filter

- What are $\mathbf{P}_{k|k}$ and $\mathbf{P}_{k|k-1}$

$$\begin{aligned}
\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1}) \\
\mathbf{P}_{k|k-1} &= \operatorname{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \\
\mathbf{P}_{k|k} &= \operatorname{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})
\end{aligned}$$

- Why do we need them? The goal is to $\min_{\mathbf{L}_k} \mathrm{E}\left[\left\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right\|^2\right]$. We should have $\dfrac{\partial\, \mathrm{E}\left[\left\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right\|^2\right]}{\partial\, \mathbf{L}_k} = \mathbf{0}$ and solve for $\mathbf{L}_k$ as $\mathbf{L}_k^*$, but it is hard to work on it directly. Instead people find out

$$\mathrm{E}\left[\left\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right\|^2\right] = \operatorname{tr}(\mathbf{P}_{k|k})$$

Now we only need to solve $\dfrac{\partial \operatorname{tr}(\mathbf{P}_{k|k})}{\partial\, \mathbf{L}_k} = \mathbf{0}$ and it turns out to be a very clear way to avoid both expectation and complex matrix derivative.

Trace and Determinant

Trace

$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$

$\mathrm{tr}(A) = \sum_i \lambda_i$

$\mathrm{tr}(A^T) = \mathrm{tr}\,(A)$

$\mathrm{tr}(\alpha A) = \alpha \cdot \mathrm{tr}(A)$

$\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$

$\mathrm{tr}(AB) \neq \mathrm{tr}(A)\,\mathrm{tr}(B)$

$\mathrm{tr}(AB) = \mathrm{tr}(BA)$

$\mathrm{tr}\left(M^{-1}AM\right) = \mathrm{tr}(A)$

Determinant

$\det(A) = \sum_{p_i \in P_n} \sigma(p_i) \prod_{j=1}^{n} a_{jp_i(j)}$

$\det(A) = \prod_i \lambda_i$

$\det(A^T) = \det(A)$

$\det(\alpha A) = \alpha^n \det(A)$

$\det(A+B) \neq \det(A) + \det(B)$

$\det(AB) = \det(A)\det(B)$

$\det(AB) = \det(BA)$

$\det\left(M^{-1}AM\right) = \det(A)$

$$\frac{\partial\,\mathrm{tr}(\mathbf{AX})}{\partial \mathbf{X}} = \frac{\partial\,\mathrm{tr}(\mathbf{XA})}{\partial \mathbf{X}} = \mathbf{A}^\top, \frac{\partial\,\mathrm{tr}\left(\mathbf{AX}^\top\right)}{\partial \mathbf{X}} = \frac{\partial\,\mathrm{tr}\left(\mathbf{X}^\top\mathbf{A}\right)}{\partial \mathbf{X}} = \mathbf{A}$$

$$\frac{\partial\,\mathrm{tr}\left(\mathbf{X}^\top\mathbf{SX}\right)}{\partial \mathbf{X}} = 2\mathbf{SX}, \frac{\partial\,\mathrm{tr}\left(\mathbf{XSX}^\top\right)}{\partial \mathbf{X}} = 2\mathbf{XS}, \text{ where } \mathbf{S} \text{ is symmetric.}$$

Expectation, Covariance Matrix, and Trace

$$\mathrm{E}\left[\left\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right\|^2\right] = \mathrm{tr}(\mathbf{P}_{k|k})$$

where $\mathbf{P}_{k|k} = \mathrm{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})$
Prove: Let $\mathbf{e} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k}$ and $\mathrm{E}[\mathbf{e}] = \bar{\mathbf{e}} = \mathbf{0}$

$$\mathrm{tr}(\mathbf{P}_{k|k}) = \mathrm{tr}(\mathrm{cov}(\mathbf{e})) = \sum_i \mathrm{E}[(e_i - \bar{e}_i)^2] = \sum_i \mathrm{E}[e_i^2 - \bar{e}_i^2] \quad \triangleleft \text{ this step is always true}$$

$$= \sum_i \mathrm{E}[e_i^2] - \sum_i \bar{e}_i^2 = \sum_i \mathrm{E}[e_i^2] - \bar{\mathbf{e}} = \sum_i \mathrm{E}[e_i^2] \quad \triangleleft\ \bar{e}_i \text{ is a constant}$$

$$= \mathrm{E}\left[\sum_i e_i^2\right] = \mathrm{E}[\|\mathbf{e}\|^2] = \mathrm{E}\left[\left\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right\|^2\right]$$

Express $P_{k|k}$ with $\mathbf{L}_k$

If we want to compute $\dfrac{\partial\, \mathrm{E}\left[\left\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right\|^2\right]}{\partial\, \mathbf{L}_k} = \dfrac{\partial\, \mathrm{tr}(\mathbf{P}_{k|k})}{\partial\, \mathbf{L}_k} = \mathbf{0}$, first need to express $P_{k|k}$ with $\mathbf{L}_k$.
Recall $\mathbf{P}_{k|k} = \mathrm{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})$ and $\mathbf{P}_{k|k-1} = \mathrm{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})$

$$\begin{aligned}
\mathbf{P}_{k|k} &= \mathrm{cov}\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right) \\
&= \mathrm{cov}\left(\mathbf{x}_k - \left[\hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k\left(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1}\right)\right]\right) \\
&= \mathrm{cov}\left(\mathbf{x}_k - \left[\hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k\left(\mathbf{C}\mathbf{x}_k + \mathbf{v}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1}\right)\right]\right) \\
&= \mathrm{cov}\left[\left(\mathbf{I} - \mathbf{L}_k\mathbf{C}\right)\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\right) - \mathbf{L}_k\mathbf{v}_k\right]
\end{aligned}$$

Since the measurement error $\mathbf{v}_k$ is uncorrelated with the other terms, this becomes
$\mathbf{P}_{k|k} = \mathrm{cov}\left[(\mathbf{I} - \mathbf{L}_k\mathbf{C})\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\right)\right] + \mathrm{cov}\left[\mathbf{L}_k\mathbf{v}_k\right]$
By the properties of vector covariance this becomes

$$\begin{aligned}
\mathbf{P}_{k|k} &= \left(\mathbf{I} - \mathbf{L}_k\mathbf{C}\right)\mathrm{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})\left(\mathbf{I} - \mathbf{L}_k\mathbf{C}\right)^{\mathrm{T}} + \mathbf{L}_k\,\mathrm{cov}\left(\mathbf{v}_k\right)\mathbf{L}_k^{\mathrm{T}} \\
&= (\mathbf{I} - \mathbf{L}_k\mathbf{C})\mathbf{P}_{k|k-1}(\mathbf{I} - \mathbf{L}_k\mathbf{C})^{\mathrm{T}} + \mathbf{L}_k\mathbf{V}\mathbf{L}_k^{\mathrm{T}} \\
&= \mathbf{P}_{k|k-1} - \mathbf{L}_k\mathbf{C}\mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}}\mathbf{L}_k^{\mathrm{T}} + \mathbf{L}_k(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}} + \mathbf{V})\mathbf{L}_k^{\mathrm{T}}
\end{aligned}$$

Optimal Kalman Gain (Equation C-1 and C-3)

$$\mathbf{L}_k^* = \min_{\mathbf{L}_k} \mathrm{E}\left[\left\|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\right\|^2\right] = \min_{\mathbf{L}_k} \mathrm{tr}(\mathbf{P}_{k|k})$$

Let

$$\frac{\partial\, \mathrm{tr}(\mathbf{P}_{k|k})}{\partial\, \mathbf{L}_k} = -2\mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}} + 2\mathbf{L}_k(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}} + \mathbf{V}) = 0$$

$$\mathbf{L}_k^* = \mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}}(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}} + \mathbf{V})^{-1} \quad \text{(C-1)}$$

Let us plug $\mathbf{L}_k^*$ into $\mathbf{P}_{k|k}$. Since
$\mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}}\mathbf{L}_k^{*\mathrm{T}} - \mathbf{L}_k^*(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}} + \mathbf{V})\mathbf{L}_k^{*\mathrm{T}} = (\mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}} - \mathbf{L}_k^*(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^{\mathrm{T}} + \mathbf{V}))\mathbf{L}_k^{*\mathrm{T}} = \mathbf{0}$,
therefore, when applying the Optimal Kalman Gain (omit $*$ from now), we have

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{L}_k\mathbf{C}\mathbf{P}_{k|k-1} = (\mathbf{I} - \mathbf{L}_k\mathbf{C})\mathbf{P}_{k|k-1} \quad \text{(C-3)}$$

Still Need to Compute $\mathbf{P}_{k|k-1}$ (P-2)

$$\begin{aligned}
\mathbf{x}_k &= \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \\
\hat{\mathbf{x}}_{k|k-1} &= \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}\mathbf{u}_k \\
\mathbf{P}_{k|k} &= \operatorname{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}) \\
\mathbf{P}_{k|k-1} &= \operatorname{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}_{k|k-1} &= \operatorname{cov}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \\
&= \operatorname{cov}(\mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k - \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} - \mathbf{B}\mathbf{u}_k) \\
&= \operatorname{cov}(\mathbf{A}(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}) + \mathbf{w}_k) \\
&= \operatorname{cov}(\mathbf{A}(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1})) + \operatorname{cov}(\mathbf{w}_k) \\
&= \mathbf{A}\mathbf{P}_{k-1|k-1}\mathbf{A}^{\mathrm{T}} + \mathbf{W}
\end{aligned}$$

Summary: "Hi-Five" Kalman Filter Algorithm

Initiate with $\hat{\mathbf{x}}_{k-1|k-1}$ and $\mathbf{P}_{k-1|k-1}$

Predict
(P-1) Predict the state
$\hat{\mathbf{x}}_{k\|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1\|k-1} + \mathbf{B}\mathbf{u}_k$
(P-2) Predict the error covariance
$\mathbf{P}_{k\|k-1} = \mathbf{A}\mathbf{P}_{k-1\|k-1}\mathbf{A}^{\mathrm{T}} + \mathbf{W}$

Take $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$

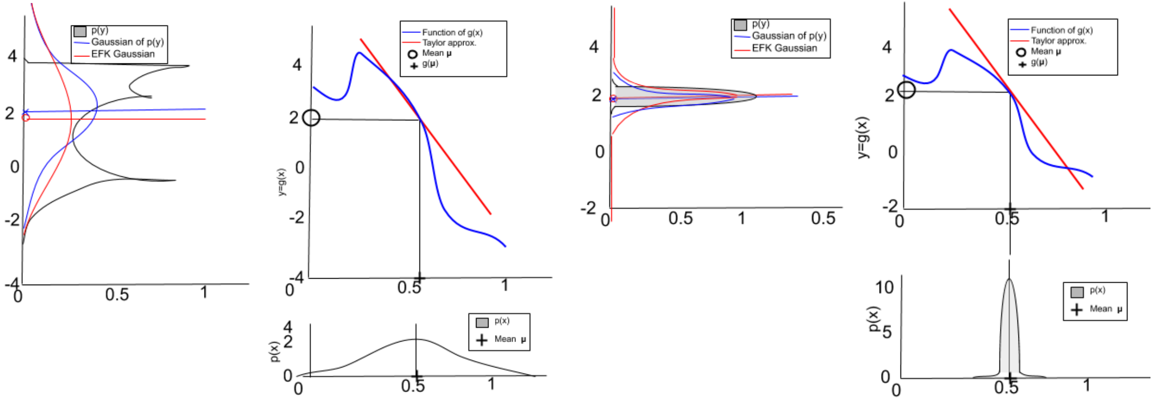
Correct
(C-1) Compute the Kalman gain
$\mathbf{L}_k = \mathbf{P}_{k\|k-1}\mathbf{C}^{\mathrm{T}}(\mathbf{C}\mathbf{P}_{k\|k-1}\mathbf{C}^{\mathrm{T}} + \mathbf{V})^{-1}$
(C-2) Update estimate with new measurement
$\hat{\mathbf{x}}_{k\|k} = \hat{\mathbf{x}}_{k\|k-1} + \mathbf{L}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_{k\|k-1})$
(C-3) Update the error covariance
$\mathbf{P}_{k\|k} = (\mathbf{I} - \mathbf{L}_k\mathbf{C})\mathbf{P}_{k\|k-1}$

How to Know $\mathbf{W}$ and $\mathbf{V}$

- In the actual implementation of the filter, the measurement noise covariance $\mathbf{V}$ is usually measured prior to operation of the filter.
- The process (system) noise covariance $\mathbf{W}$ is generally more difficult to estimate as we typically do not have the ability to directly observe the process. If the measurement noise is small, we may choose a small $\mathbf{W}$ to force the output behave more certain and measure the randomness.
- In either case, whether or not we have a rational basis for choosing the parameters, often times superior filter performance (statistically speaking) can be obtained by tuning the filter parameters $\mathbf{W}$ and $\mathbf{V}$.
- $\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{L}_k\mathbf{C})\mathbf{P}_{k|k-1}$ is computationally cheaper and thus nearly always used in practice, but is only correct for the optimal gain. If arithmetic precision is unusually low causing problems with numerical stability, this simplification cannot be applied; Joseph form must be used.

Table of Contents

Assumption violated for Nonlinear System



If $p(x)$ has high variance relative to region in which linearization is inaccurate.

If $p(x)$ has small variance relative to region in which linearization is accurate

Extended Kalman Filter - A Direct Intuition using Linearization

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k) + \mathbf{w}_k$$
$$\mathbf{y}_k = h(\mathbf{x}_k) + \mathbf{v}_k$$

Let $\mathbf{F}_k = \left.\frac{\partial f}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_k}, \mathbf{H}_k = \left.\frac{\partial h}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_{k|k-1}}$ Replace $\mathbf{A}$ with $\mathbf{F}_k$, $\mathbf{C}$ with $\mathbf{H}_k$. We get EKF.

Initiate with $\hat{\mathbf{x}}_{k-1|k-1}$ and $\mathbf{P}_{k-1|k-1}$

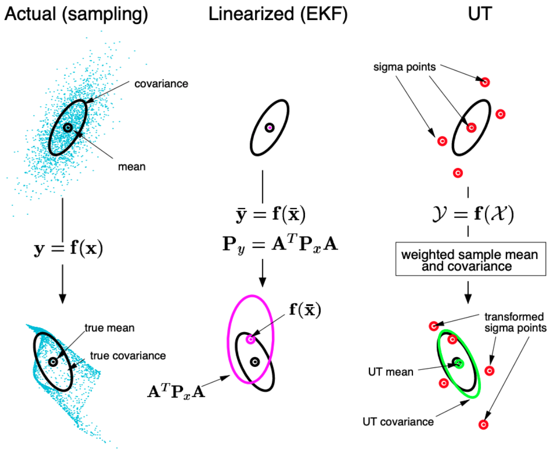
Predict
(P-1) Predict the state
$\hat{\mathbf{x}}_{k\|k-1} = f(\hat{\mathbf{x}}_{k-1\|k-1}, \mathbf{u}_k)$
(P-2) Predict the error covariance
$\mathbf{P}_{k\|k-1} = \mathbf{F}_k \mathbf{P}_{k-1\|k-1} \mathbf{F}_k^{\mathrm{T}} + \mathbf{W}_k$

Take $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$

Correct
(C-1) Compute the Kalman gain
$\mathbf{L}_k = \mathbf{P}_{k\|k-1} \mathbf{H}_k^{\mathrm{T}} (\mathbf{H}_k \mathbf{P}_{k\|k-1} \mathbf{H}_k^{\mathrm{T}} + \mathbf{V}_k)^{-1}$
(C-2) Update estimate with new measurement
$\hat{\mathbf{x}}_{k\|k} = \hat{\mathbf{x}}_{k\|k-1} + \mathbf{L}_k(\mathbf{y}_k - h(\hat{\mathbf{x}}_{k\|k-1}))$
(C-3) Update the error covariance
$\mathbf{P}_{k\|k} = (\mathbf{I} - \mathbf{L}_k \mathbf{H}_k) \mathbf{P}_{k\|k-1}$

Table of Contents

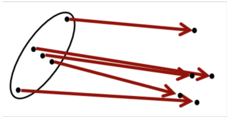
Issues of EKF for "Very" Nonlinear System



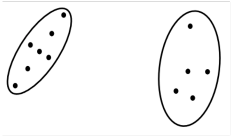
Courtesy: E. A. Wan and R. van der Merwe

Concept of the Unscented Kalman Filter

1. Compute a set of (so-called) sigma points. Each sigma points has a weight.



2. Transform each sigma point through the non-linear function



3. Compute Gaussian from the transformed and weighted points

Courtesy: Cyrill Stachniss

Sigma Points

For a random vector $\mathbf{x} = (x_1, \ldots, x_L)$, sigma points are any set of vectors

$$\{\mathbf{s}_0, \ldots, \mathbf{s}_N\} = \left\{ \begin{pmatrix} s_{0,1} & s_{0,2} & \ldots & s_{0,L} \end{pmatrix}, \ldots, \begin{pmatrix} s_{N,1} & s_{N,2} & \cdots & s_{N,L} \end{pmatrix} \right\}$$

attributed with

- first-order weights $W_0^a, \ldots, W_N^a$ that fulfill
  1. $\sum_{j=0}^{N} W_j^a = 1$
  2. $E[x_i] = \sum_{j=0}^{N} W_j^a s_{j,i}$ for all $i = 1, \ldots, L$
- second-order weights $W_0^c, \ldots, W_N^c$ that fulfill
  1. $\sum_{j=0}^{N} W_j^c = 1$
  2. $E[x_i x_l] = \sum_{j=0}^{N} W_j^c s_{j,i} s_{j,l}$ for all pairs $(i, l) \in \{1, \ldots, L\}^2$

One Popular Choice of Sigma Points

Sigma points	First-order (mean) Weights	Second-order (co-variance) weights
$\mathbf{s}_0 = \hat{\mathbf{x}}_{k-1\|k-1}$	$W_0^a = \frac{\alpha^2\kappa - L}{\alpha^2\kappa}$	$W_0^c = W_0^a + 1 - \alpha^2 + \beta$
$\mathbf{s}_j = \hat{\mathbf{x}}_{k-1\|k-1} + \alpha\sqrt{\kappa}\mathbf{A}_j$ $\mathbf{s}_{L+j} = \hat{\mathbf{x}}_{k-1\|k-1} - \alpha\sqrt{\kappa}\mathbf{A}_j$ $j = 1, \dots, L$	$W_j^a = \dfrac{1}{2\alpha^2\kappa}$ $j = 1, \dots, 2L$	$W_j^c = W_j^a$

Choice of hyper-parameters

- α and κ control the spread of the signma points. Usually $\alpha = 10^{-3}, \kappa = 1$ is recommended
- β is related to the distribution of x. If the true distribution of x is Gaussian, $\beta = 2$ is optimal.
- The $\mathbf{A}$ where $\mathbf{P}_{k-1|k-1} = \mathbf{A}\mathbf{A}^\mathsf{T}$. The matrix $\mathbf{A}$ should be calculated using numerically efficient and stable methods such as the Cholesky decomposition.

Predict

- Given the estimates of the mean and covariance $\hat{\mathbf{x}}_{k-1|k-1}$ and $\mathbf{P}_{k-1|k-1}$, we can obtains $N = 2L + 1$ sigma points. The sigma points are propagated through the transition function

$$\mathbf{x}_j = f\left(\mathbf{s}_j\right) \quad j = 0, \ldots, 2L$$

- The propagated sigma points are weighed to produce the predicted mean and covariance

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{j=0}^{2L} W_j^a \mathbf{x}_j$$

$$\mathbf{P}_{k|k-1} = \sum_{j=0}^{2L} W_j^c \left(\mathbf{x}_j - \hat{\mathbf{x}}_{k|k-1}\right)\left(\mathbf{x}_j - \hat{\mathbf{x}}_{k|k-1}\right)^\top + \mathbf{W}_k$$

Algorithms of the Unscented Kalman Filter (continued)

Correct

- Given prediction estimates $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$, a new set of $N = 2L + 1$ sigma points $\mathbf{s}_0, \ldots, \mathbf{s}_{2L}$ with corresponding weights is calculated. The sigma points are transformed through measurement function

$$\mathbf{z}_j = h\left(\mathbf{s}_j\right), j = 0, 1, \ldots, 2L$$

Then the empirical mean and covariance of the transformed points are calculated

$$\hat{\mathbf{z}} = \sum_{j=0}^{2L} W_j^a \mathbf{z}_j$$

$$\hat{\mathbf{S}}_k = \sum_{j=0}^{2L} W_j^c \left(\mathbf{z}_j - \hat{\mathbf{z}}\right)\left(\mathbf{z}_j - \hat{\mathbf{z}}\right)^\top + \mathbf{V}_k$$

Algorithms of the Unscented Kalman Filter (continued)

Correct (continued)

- The cross covariance matrix is also needed

$$\mathbf{C}_{\mathbf{sz}} = \sum_{j=0}^{2L} W_j^c \left(\mathbf{s}_j - \hat{\mathbf{x}}_{k|k-1}\right)\left(\mathbf{z}_j - \hat{\mathbf{z}}\right)^\top$$

  where $\mathbf{s}_j$ are the untransformed sigma points created from $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$.
- The Kalman gain can be calculated as

$$\mathbf{L}_k = \mathbf{C}_{\text{sz}} \hat{\mathbf{S}}_k^{-1}$$

- The corrected mean and variance estimates are

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k \left(\mathbf{y}_k - \hat{\mathbf{z}}\right)$$
$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{L}_k \hat{\mathbf{S}}_k \mathbf{L}_k^\top$$

EKF VS UKF

EKF: $\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k) + \mathbf{w}_k$
$\mathbf{y}_k = h(\mathbf{x}_k) + \mathbf{v}_k$

$$\mathbf{F}_k = \left.\frac{\partial f}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_{k-1|k-1},\mathbf{u}_k}, \mathbf{H}_k = \left.\frac{\partial h}{\partial \mathbf{x}}\right|_{\hat{\mathbf{x}}_{k|k-1}}$$

Initiate with $\hat{\mathbf{x}}_{k-1|k-1}$ and $\mathbf{P}_{k-1|k-1}$

Predict

(P-1) Predict the state

$\hat{\mathbf{x}}_{k|k-1} = f(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_k)$

(P-2) Predict the error covariance

$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^{\mathrm{T}} + \mathbf{W}_k$

Take $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$

Correct

(C-1) Compute the Kalman gain

$\mathbf{L}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^{\mathrm{T}}(\mathbf{H}_k \mathbf{P}_{k|k-1}\mathbf{H}_k^{\mathrm{T}} + \mathbf{V}_k)^{-1}$

(C-2) Update estimate with new measurement

$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k(\mathbf{y}_k - h(\hat{\mathbf{x}}_{k|k-1}))$

(C-3) Update the error covariance

$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{L}_k \mathbf{H}_k)\mathbf{P}_{k|k-1}$

Initiate with $\hat{\mathbf{x}}_{k-1|k-1}$ and $\mathbf{P}_{k-1|k-1}$

Predict

(P-1) Predict the state

$\mathbf{s}_0 = \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{s}_j = \hat{\mathbf{x}}_{k-1|k-1} + \alpha\sqrt{\kappa}\mathbf{A}_j, \mathbf{A}\mathbf{A}^{\intercal} = \mathbf{P}_{k-1|k-1}$

$\mathbf{s}_{L+j} = \hat{\mathbf{x}}_{k-1|k-1} - \alpha\sqrt{\kappa}\mathbf{A}_j, \quad j = 1, \ldots, L$

$\hat{\mathbf{x}}_{k|k-1} = \sum_{j=0}^{2L} W_j^a \mathbf{x}_j, \mathbf{x}_j = f(\mathbf{s}_j, \mathbf{u}_k), \quad j = 0, \ldots, 2L$

$W_0^a = \frac{\alpha^2\kappa - L}{\alpha^2\kappa}, W_j^a = \frac{1}{2\alpha^2\kappa}$

(P-2) Predict the error covariance

$\mathbf{P}_{k|k-1} = \sum_{j=0}^{2L} W_j^c \left(\mathbf{x}_j - \hat{\mathbf{x}}_{k|k-1}\right)\left(\mathbf{x}_j - \hat{\mathbf{x}}_{k|k-1}\right)^{\top} + \mathbf{W}_k$

$W_0^c = W_0^a + 1 - \alpha^2 + \beta, W_j^c = W_j^a$

Take $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$

Correct

(C-1) Compute the Kalman gain

$\mathbf{L}_k = \mathbf{C}_{\mathrm{sz}}\hat{\mathbf{S}}_k^{-1}, \mathbf{z}_j = h(\mathbf{s}_j), j = 0, 1, \ldots, 2L, \hat{\mathbf{z}} = \sum_{j=0}^{2L} W_j^a \mathbf{z}_j$

$\hat{\mathbf{S}}_k = \sum_{j=0}^{2L} W_j^c (\mathbf{z}_j - \hat{\mathbf{z}})(\mathbf{z}_j - \hat{\mathbf{z}})^{\top} + \mathbf{V}_k$

$\mathbf{C}_{\mathbf{sz}} = \sum_{j=0}^{2L} W_j^c \left(\mathbf{s}_j - \hat{\mathbf{x}}_{k|k-1}\right)(\mathbf{z}_j - \hat{\mathbf{z}})^{\top}$

(C-2) Update estimate with new measurement

$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{L}_k(\mathbf{y}_k - \hat{\mathbf{z}})$

(C-3) Update the error covariance $\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{L}_k\hat{\mathbf{S}}_k\mathbf{L}_k^{\top}$

Table of Contents

Definition of SLAM

- Building a map and locating the robot in the map at the same time
- Given
  - Robots control input $u_{1:T} = \{u_1, u_2, u_3, \ldots, u_T\}$
  - Sensor measurements $y_{1:T} = \{y_1, y_2, y_3, \ldots, y_T\}$
- Want
  - Map of the environment m
  - Trajectory of the robot $x_{0:T} = \{x_0, x_1, x_2, \ldots, x_T\}$

Apply EKF to SLAM

- Estimate robot's pose and location of map features in the environment.
- Assume we know the correspondence of map features across different time steps.
- State space is $x_t = [\underbrace{X_t, Y_t, \psi_t}_{\text{robot's pose}}, \underbrace{m_x^1, m_y^1}_{\text{landmark 1}}, \cdots, \underbrace{m_x^n, m_y^n}_{\text{landmark n}}]^T$, where X_t, Y_t, ψ_t are the global position and heading angle of the robot. m_x^j and m_y^j are the global position of the map feature j for $j = 1, \cdots, n$.
- The dynamical system of the robot is

$$
\begin{aligned}
X_{t+1} &= X_t + \delta t \dot{X}_t + \omega_t^x \\
Y_{t+1} &= Y_t + \delta t \dot{Y}_t + \omega_t^y \\
\psi_{t+1} &= \psi_t + \delta t \dot{\psi}_t + \omega_t^\psi
\end{aligned}
$$

where δt is the time step, and ω_t^x, ω_t^y, ω_t^ψ are the dynamical noise.

Apply EKF to SLAM (continued)

- A robot will have both range and bearing measurements relative to the map features (landmarks), which can be achieved by using sensors such as LiDAR.
- The range measurement is defined as the distance to each feature with the measurement equations $y^j_{t,distance} = \left\| m^j - p_t \right\| + v^j_{t,distance}$ for $j = 1, ..., n$, where $m^j = [m^j_x, m^j_y]^T$, $p_t = [X_t, Y_t]^T$ and $v^j_{t,distance}$ is the range sensor measurement noise.
- The bearing measure is defined as the angle between the vehicle's heading (yaw angle) and ray from the vehicle to the feature with the measurement equations $y^j_{t,bearing} = atan2(m^j_y - Y_t, m^j_x - X_t) - \psi_t + v^j_{t,bearing}$ for $j = 1, ..., n$, where $v^j_{t,bearing}$ is the bearing sensor measurement noise.
- We can apply EKF to estimate the state including the robot's position, heading angle and the location of the map features.
- You will derive matrices A, B, C, D for EKF SLAM in your project.