Extended Kalman Filter

Extended Kalman Filter

Extended Kalman Filters keep the first assumption

1. Everything is modeled as multivariate Gaussians,

But relax the second assumption

2. All models are **locally** linear

Transition and measurement models

$$\begin{aligned} x_t = Ax_{t-1} + Bu_t + \epsilon_x &\rightarrow \quad g(u_t, x_{t-1}) + \epsilon_x \\ z_t = Cx_t + \epsilon_z &\rightarrow \quad h(x_t) + \epsilon_z \end{aligned}$$

Linearization

First-order Taylor Series

$$f(x) \approx f(a) + \frac{\partial f(a)}{\partial x}(x - a)$$

Vector case

For $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$

$$\begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{bmatrix} \approx \begin{bmatrix} f_1(\mathbf{a}) \\ f_2(\mathbf{a}) \\ \vdots \\ f_n(\mathbf{a}) \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(\mathbf{x})}{\partial x_1} & \frac{\partial f_n(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_n(\mathbf{x})}{\partial x_n} \end{bmatrix} \begin{bmatrix} x_1 - a_1 \\ x_2 - a_2 \\ \vdots \\ x_n - a_n \end{bmatrix}$$

Linearization around local point

Using the Taylor expansion

Prediction:

$$
\begin{aligned}
g(u_t, x_{t-1}) &\approx g(u_t, \mu_{t-1}) + \frac{\partial g(u_t, \mu_{t-1})}{\partial x_{t-1}}(x_{t-1} - \mu_{t-1}) \\
&\approx G x_{t-1} - G \mu_{t-1} + g(u_t, \mu_{t-1})
\end{aligned}
$$

where

$$
G = \frac{\partial g(u_t, \mu_{t-1})}{\partial x_{t-1}}
$$

Linearization around local point

Using the Taylor expansion

Correction

$$
\begin{aligned}
h(x_t) &\approx h(\bar{u}_t) + \frac{\partial h(\bar{\mu}_t)}{\partial x_t}(x_t - \bar{\mu}_t) \\
&\approx Hx_t - H\bar{\mu}_t + h(\bar{\mu}_t)
\end{aligned}
$$

where

$$
H = \frac{\partial h(\bar{\mu}_t)}{\partial x_t}
$$

Extended Kalman Filter

Prediction Step

1. $\bar{\mu}_t = g(u_t, \mu_{t-1})$
2. $\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T + R$

Correction Step

3. $K_t = \bar{\Sigma}_t {H_t}^T (Q + H_t \bar{\Sigma}_t H_t^T)^{-1}$
4. $\mu_t = \bar{\mu}_t + K_t(z_t - h(\bar{\mu}_t))$
5. $\Sigma_t = (I - K_t H_t)\bar{\Sigma}_t$

Understanding Kalman Filters: Perfect Motion

Perfect motion would translate to $R \to [\mathbf{0}]$

$$\bar{\mu}_t = g(u_t, \mu_{t-1})$$

$$\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T$$

We can see that the updated covariance matrix is the direct result of the linear transformation applied to the previous state.

Understanding Kalman Filters: Infinite Variance Motion

Infinite variance motion would translate to $R \to \infty$

$$\bar{\mu}_t = g(u_t, \mu_{t-1})$$
$$\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T + \infty = \infty$$

The covariance matrix tends to infinity which means that our guess on the robot's location is now uninformative

Understanding Kalman Filters: Perfect Sensor model

Perfect motion would translate to $Q \rightarrow [\mathbf{0}]$

$$\begin{aligned} K_t &= \bar{\Sigma}_t H_t^T([\mathbf{0}] + H_t \bar{\Sigma}_t H_t^T)^{-1} \\ &= \bar{\Sigma}_t H_t^T (H_t^T)^{-1} \bar{\Sigma}_t^{-1} H_t^{-1} \\ &= H_t^{-1} \end{aligned}$$

Mean

For the KF case, the mean is computed as

$$\begin{aligned}\mu_t &= \bar{\mu}_t + C_t^{-1}(z_t - C(\bar{\mu}_t)) & = \bar{\mu}_t + C_t^{-1} z_t - \bar{\mu}_t \\ &= C_t^{-1} z_t\end{aligned}$$

That is, the new mean completely tosses out previous state $\bar{\mu}_t$ and completely believes on the location generated by the measurement.

`Note:` Remember that for the linear case z

For the EKF, the mean is computed as

$$\mu_t = \bar{\mu}_t + H_t^{-1}(z_t - h(\bar{\mu}_t))$$

Therefore, it also assigns μ_t as the belief on the location generated by the measurement but considering an additional term to correct for the non-linearity of the system

Variance

For this case, we have that the variance

$$\Sigma_t = (I - H_t^{-1} H_t)\bar{\Sigma}_t = [\mathbf{0}]$$

For this case, the new covariance matrix becomes zero, as we are certain that the measurement is correct.

Understanding Kalman Filters: Infinite Variance Sensor

A sensor that provides no additional information, or is unreliable would have $Q \to \infty$

$$K_t = \bar{\Sigma}_t H_t^T (\infty + H_t \bar{\Sigma}_t H_t^T)^{-1} = [\mathbf{0}]$$
$$\mu_t = \bar{\mu}_t$$
$$\Sigma = \bar{\Sigma}$$

The update state is not performed. That is, the sensor measurement is ignored.

Main problems with KF/EKF

1. EKF will only work if the model is more-less linearly local at the belief location.
    - If the function is too non-linear, it fails
    - The further our belief is from the true location, the worse the system will perform. An error in prediction can escalate to localization failure
2. It is difficult for the EKF to solve the **Global Localization Problem**
    - A variant called the Multi-Hypothesis Tracker can address this issue

Example Application

Localization using EKF using the odometry model studied in class 4, and the measurement model for landmarks studied in class 6.

Motion: Odometry model.

For $u = \begin{bmatrix}\delta_{rot1} & \delta_{trans} & \delta_{rot1}\end{bmatrix}^T$

$$\begin{bmatrix} x' \\ y' \\ \theta' \end{bmatrix} = \begin{bmatrix} x + \delta_{trans}\cos(\theta + \delta_{rot1}) \\ y + \delta_{trans}\sin(\theta + \delta_{rot1}) \\ \theta + \delta_{rot1} + \delta_{rot2} \end{bmatrix} + \mathcal{N}(0, R)$$

Motion model Linearization

$$g(u_t, x_{t-1}) \approx g(u_t, u_{t-1}) + G_t(x_{t-1} - \mu_{t-1})$$

$$G_t = \begin{bmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} & \frac{\partial x'}{\partial \theta} \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} & \frac{\partial y'}{\partial \theta} \\ \frac{\partial \theta'}{\partial x} & \frac{\partial \theta'}{\partial y} & \frac{\partial \theta'}{\partial \theta} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & -\delta_{trans} \sin(\theta + \delta_{rot1}) \\ 0 & 1 & \delta_{trans} \cos(\theta + \delta_{rot1}) \\ 0 & 0 & 1 \end{bmatrix}$$

What about motion noise?

From our motion model, the noise is assumed to be Gaussian with respect to
$\delta_{rot1,trans,rot2}$
But, from the motion model, $g(u_t, x_{t-1})$ is also nonlinear with respect to them, so we need to linearize it

Same as for G_t, we compute V_t as

$$V_t = \begin{bmatrix} \frac{\partial x'}{\partial \delta_{rot1}} & \frac{\partial x'}{\partial \delta_{trans}} & \frac{\partial x'}{\partial \delta_{rot2}} \\ \frac{\partial y'}{\partial \delta_{rot1}} & \frac{\partial y'}{\partial \delta_{trans}} & \frac{\partial y'}{\partial \delta_{rot2}} \\ \frac{\partial \theta'}{\partial \delta_{rot1}} & \frac{\partial \theta'}{\partial \delta_{trans}} & \frac{\partial \theta'}{\partial \delta_{rot2}} \end{bmatrix}$$
$$= \begin{bmatrix} -\delta_{trans}\sin(\theta + \delta_{rot1}) & \cos(\theta + \delta_{rot1}) & 0 \\ \delta_{trans}\cos(\theta + \delta_{rot1}) & \sin(\theta + \delta_{rot1}) & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Observation model

Assuming we have a landmark-based system, where we get both range and bearing.

For the i-th feature, we get

$$z_t^i = \begin{bmatrix} r_t^i \\ \phi_t^i \end{bmatrix} = \begin{bmatrix} \sqrt{(m_{j,x} - x)^2 + (m_{j,y} - y)^2} \\ atan2(m_{j,y} - y, m_{j,x} - x) - \theta \end{bmatrix} + \mathcal{N}(0, Q)$$

Linearization

$$H = \begin{bmatrix} -\frac{m_{j,x} - \bar{\mu}_{t,x}}{\sqrt{q}} & -\frac{m_{j,y} - \bar{\mu}_{t,y}}{\sqrt{q}} & 0 \\ \frac{m_{j,y} - \bar{\mu}_{t,y}}{q} & -\frac{m_{j,x} - \bar{\mu}_{t,x}}{q} & -1 \end{bmatrix}$$

where $q = (m_{j,x} - \bar{\mu}_{t,x})^2 + (m_{j,y} - \bar{\mu}_{t,y})^2$

Extended Kalman Filter

Prediction Step

$$\bar{\mu}_t = g(u_t, \mu_{t-1})$$

$$R = V_t \begin{bmatrix} \alpha_1 \delta_{rot1}^2 + \alpha_2 \delta_{trans}^2 & 0 & 0 \\ 0 & \alpha_3 \delta_{trans}^2 + \alpha_4 (\delta_{rot1}^2 + \delta_{rot2}^2) & 0 \\ 0 & 0 & \alpha_1 \delta_{rot1}^2 + \alpha_2 \delta_{trans}^2 \end{bmatrix} V_t^T$$

$$\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T + R$$

Correction Step

$$Q = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\phi^2 \end{bmatrix}$$

for all z_i **do**

$$K_t = \bar{\Sigma}_t H_t^T (Q + H_t \bar{\Sigma}_t H_t^T)^{-1}$$
$$\bar{\mu}_t = \bar{\mu}_t + K_t(z_t - h(\bar{\mu}_t))$$
$$\bar{\Sigma}_t = (I - K_t H_t)\bar{\Sigma}_t$$

Output

$$\mu_t = \bar{\mu}_t$$
$$\Sigma_t = \bar{\Sigma}_t$$