Extended Kalman Filter

The Kalman filter is derived for a Linear Gaussian system. For nonlinear system, one can apply linear approximation → extended Kalman Filter:

Consider nonlinear filtering problem

$$X_{t+1} = m(X_t) + \varepsilon_t^m, \quad \varepsilon_t^m \sim N(0, Q)$$
$$y_{t+1}^o = h(X_{t+1}) + \varepsilon_{t+1}^o, \quad \varepsilon_{t+1}^o \sim N(0, R)$$

Assume m and h can be approximated by

$$m(X_t) \cong m(\bar{X}_t^a) + \left.\frac{\partial m}{\partial x}\right|_{\bar{X}_t^a} (X_t - \bar{X}_t^a)$$

$$h(X_{t+1}) \cong h(\bar{X}_{t+1}^b) + \left.\frac{\partial h}{\partial x}\right|_{\bar{X}_{t+1}^b} (X_{t+1} - \bar{X}_{t+1}^b)$$

Denote $M_t \equiv \left.\frac{\partial m}{\partial x}\right|_{\bar{X}_t^a}$, $H_{t+1} \equiv \left.\frac{\partial h}{\partial x}\right|_{\bar{X}_{t+1}^b}$

→ Now M and H operators depend on $\bar{x}$ values, update eqns using M_t, H_{t+1} are not offline anymore!

We have

$$X_{t+1} = m(\bar{X}_t^a) + M_t(X_t - \bar{X}_t^a) + \varepsilon_t^m$$

$$y_{t+1}^o = h(\bar{X}_{t+1}^b) + H_{t+1}(X_{t+1} - \bar{X}_{t+1}^b) + \varepsilon_{t+1}^o$$

$$\bar{x}^b_{t+1} = \mathbb{E}(x_{t+1}) = m(\bar{x}^a_t)$$

$$P^b_{t+1} = \mathbb{E}(\varepsilon^b_{t+1}\varepsilon^{bT}_{t+1})$$

$$\varepsilon^b_{t+1} = x_{t+1} - \bar{x}^b_{t+1} = M_t(x_t - \bar{x}^a_t) + \varepsilon^m_t$$
$$= M_t\varepsilon^a_t + \varepsilon^m_t$$

$$= M_t\mathbb{E}(\varepsilon^a_t\varepsilon^{aT}_t)M_t^T + \mathbb{E}(\varepsilon^m_t\varepsilon^{mT}_t)$$

$$= M_t P^a_t M_t^T + Q$$

Similar to the derivation of Kalman filter.

$$\bar{x}^a_{t+1} = \bar{x}^b_{t+1} + K_{t+1}\left(y^o_{t+1} - h(\bar{x}^b_{t+1})\right)$$

Where K_{t+1} is found so that $tr(P^a_{t+1})$ is minimum
→ best estimate.

$$\varepsilon^a_{t+1} = x_{t+1} - \bar{x}^a_{t+1} = \underbrace{x_{t+1} - \bar{x}^b_{t+1}}_{=\varepsilon^b_{t+1}} - K_{t+1}\left(H_{t+1}\varepsilon^b_{t+1} + \varepsilon^o_{t+1}\right)$$

$$= (I - K_{t+1}H_{t+1})\varepsilon^b_{t+1} - K_{t+1}\varepsilon^o_{t+1}$$

$$P^a_{t+1} = (I - K_{t+1}H_{t+1})P^b_{t+1}(I - K_{t+1}H_{t+1})^T + K_{t+1}RK^T_{t+1}$$

$$K_{t+1} = P^b_{t+1}H^T_{t+1}(H_{t+1}P^b_{t+1}H^T_{t+1} + R)^{-1}$$

1 forecast step

$$\bar{x}^b_{t+1} = m(\bar{x}^a_t) \quad (1)$$

$$P^b_{t+1} = M_t P^a_t M^T_t + Q \quad (2)$$

2. analysis step

$$\bar{x}^a_{t+1} = \bar{x}^b_{t+1} + K_{t+1}\left(y^o_{t+1} - h(\bar{x}^b_{t+1})\right) \quad (3)$$

$$P^a_{t+1} = (I - K_{t+1} H_{t+1}) P^b_{t+1} \quad (4)$$

$$K_{t+1} = P^b_{t+1} H^T_{t+1} (H_{t+1} P^b_{t+1} H^T_{t+1} + R)^{-1} \quad (5)$$

Problem: For large system, Extended KF is practically useless since propagating the full covariance matrix is very expensive.

→ Suppose it takes 6 minutes to propagate $\bar{x}^b_{t+1} = m(\bar{x}^a_t)$.
Say state dimension is $n = 3$ million.
It will take 30,000 hours to get $P^b_{t+1} = M_t P^a_t M^T_t + Q$.

→ assuming error distribution to be Gaussian ignores higher moments of error statistics.
- If nonlinear error evolution cause higher moments to be non-zero, the analysis for $\bar{x}$, P will be suboptimal.
- Need to update higher moments of error distribution, but can not find proper closure of update equations.