Notes on Kalman Filter (KF, EKF, ESKF)

Edward Gyubeom Im*

February 4, 2023

Contents

1 Probability theory

1.1 Sample space and event

어떤 시행에서 일어날 수 있는 모든 결과들의 모임을 표본공간 Ω라고 한다. 예를 들어 주사위를 한번 던지는 시행의 경우 표본공간은 $\{1, 2, 3, 4, 5, 6\}$과 같은 집합이 된다. 표본공간 Ω의 부분집합을 사건(event) $\mathcal{F}$라고 한다.

*blog: alida.tistory.com, email: gyurse@gmail.com

1.2 Probability Function

확률함수 p는 표본공간의 원소를 0과 1사이의 숫자에 대응시키는 함수를 의미한다. 사건 $\mathcal{F}$ 에 대한 확률은 다음과 같이 정의할 수 있다.

$$\forall \mathcal{F} \in \Omega, \quad p(\mathcal{F}) = \sum_{w \in \mathcal{F}} p(w) \tag{1}$$

1.3 Random variables and distribution

확률변수(random variable)는 표본공간 Ω에 정의된 함수를 의미한다. 이 때 확률변수의 결과값은 항상 실수이다. 분포(distribution)은 확률변수가 가질 수 있는 값들에 대해서 확률들을 나열해 놓은 것을 의미한다.

중요한 점은 어떤 확률변수 x, y가 확률함수 p에 대해 같은 분포를 가져도 둘은 다른 확률변수일 수 있다.

1.4 Expected value

표본공간 Ω에서 정의된 확률변수 x가 있을 때 확률함수 p에 대한 x의 기대값은 $\mathbb{E}[x]$라고 하고 다음과 같은 식으로 나타낸다.

$$\mathbb{E}[x] = \sum_{w \in \Omega} x(w)p(w) \tag{2}$$

기대값은 선형성(Linearity)라는 성질을 가지고 있다. 수학에서 선형성에 대한 정의는 다음과 같다. 임의의 함수 f에 대해

- 임의의 수 x,y에 대해 f(x+y) = f(x)+f(y)가 항상 성립하고
- 임의의 수 x와 a에 대해 f(ax) = af(x)가 항상 성립하면
- 함수 f는 선형이라고 한다.

따라서 임의의 확률변수 x, y와 임의의 실수 a, b에 대해서 다음 식이 성립하게 된다.

$$\mathbb{E}[ax + by] = a\mathbb{E}[x] + b\mathbb{E}[y] \tag{3}$$

그리고 선형인 함수 $L(x)$에 대해서 기대값과 함수의 계산순서를 바꿀 수 있다.

$$\mathbb{E}[L(x)] = L(\mathbb{E}[x]) \tag{4}$$

1.5 Variance

x의 분산은 var$[x]$라고 하고 다음과 같은 식으로 나타낸다.

$$\begin{aligned} \text{var}[x] &= \mathbb{E}[(x - \mathbb{E}(x))^2] \\ \text{var}[x] &= \mathbb{E}[x^2] - (\mathbb{E}[x])^2 \end{aligned} \tag{5}$$

1.6 Probability axioms

표본공간 Ω에 사건 $\mathcal{F}$가 있을 때, 사건 $\mathcal{F}$의 확률변수 x가 일어날 확률 $p(x)$는 항상 0 이상 1 이하이다.

$$0 \leq p(x) \leq 1 \quad \forall x \in \mathcal{F} \tag{6}$$

표본공간 Ω 전체가 일어날 확률은 1이다.

$$p(\Omega) = 1 \tag{7}$$

1.7 Gaussian distribution

스칼라 확률변수 x가 가우시안 분포를 따른다고 하면

$$x \sim \mathcal{N}(\mu, \sigma^2) \tag{8}$$

가 성립한다. 이 때, 확률분포함수 $p(x)$는 다음과 같이 정의된다.

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \tag{9}$$

1.8 Joint gaussian distribution

확률변수가 두 개 이상일 때는 다변수 확률분포(multivariate probability distribution)를 사용해야한다. 예를 들어 두 개의 확률변수 a, b가 있을 때 다변수 확률분포$p(a, b) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$는 다음과 같이 나타낼 수 있다.

$$p(a,b) = \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}} \exp -\frac{1}{2}\left(\begin{bmatrix} a - \mathbb{E}(a) \\ b - \mathbb{E}(b) \end{bmatrix} \right)^T \boldsymbol{\Sigma}^{-1} \left(\begin{bmatrix} a - \mathbb{E}(a) \\ b - \mathbb{E}(b) \end{bmatrix} \right) \tag{10}$$

이 때 평균은 $\boldsymbol{\mu} = \begin{bmatrix} \mathbb{E}(a) \\ \mathbb{E}(b) \end{bmatrix}$이고 분산은 $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb} \end{bmatrix}$이다. 이 때, 분산은 여러 변수에 대한 분산을 의미하고 대각성분들은 하나의 변수에 대한 분산을 의미하며 대각성분이 아닌 성분들은 두 변수 간 상관 관계를 의미한다. **이러한 다변수 확률분포에서 분산을 일반적으로 공분산(covariance)라고 부른다.**

1.9 Multivariate gaussian distribution

벡터 확률변수 $\mathbf{x} \in \mathbb{R}^n$가 가우시안 분포를 따른다고 하면

$$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{11}$$

가 성립한다. 평균 $\boldsymbol{\mu} \in \mathbb{R}^n$은 벡터이고 공분산 $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$은 행렬이다. 이 때, 확률분포함수 $p(\mathbf{x})$는 다음과 같이 정의된다.

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Omega} (\mathbf{x} - \boldsymbol{\mu}) \right) \tag{12}$$

$|\boldsymbol{\Sigma}|$은 $\boldsymbol{\Sigma}$의 행렬식(determinant)을 의미하며 $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$을 의미한다.

1.10 Linear transformation of gaussian random variable

스칼라 랜덤 변수(random variable) x가 주어졌을 때 만약 가우시안 분포를 따른다고 가정하면 다음과 같이 표기할 수 있다.

$$x \sim \mathcal{N}(\mu, \sigma^2) \tag{13}$$

벡터 랜덤 변수 $\mathbf{x} \in \mathbb{R}^n$가 가우시안 분포를 따를 때는 다음과 같이 표기할 수 있다.

$$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{14}$$

$\mathbf{x}$를 선형 변환(linear transformation)한 새로운 랜덤변수 $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$가 주어졌다고 하면 $\mathbf{y}$는 아래와 같은 확률 분포를 따른다.

$$\boxed{\begin{aligned} \mathbf{y} &= \mathbf{A}\mathbf{x} + \mathbf{b} \\ &\sim \mathcal{N}(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\intercal) \end{aligned}} \tag{15}$$

공분산 $\text{cov}(\mathbf{A}\mathbf{x} + \mathbf{b})$는 다음과 같이 유도할 수 있다.

$$\begin{aligned} \text{cov}(\mathbf{A}\mathbf{x} + \mathbf{b}) &= \mathbb{E}((\mathbf{y} - \boldsymbol{\mu}_\mathbf{y})(\mathbf{y} - \boldsymbol{\mu}_\mathbf{y})^\intercal) \\ &= \mathbb{E}((\mathbf{y} - (\mathbf{A}\boldsymbol{\mu} + \mathbf{b}))(\mathbf{y} - (\mathbf{A}\boldsymbol{\mu} + \mathbf{b}))^\intercal) \\ &= \mathbb{E}(((\mathbf{A}\mathbf{x} + \mathbf{b}) - (\mathbf{A}\boldsymbol{\mu} + \mathbf{b}))((\mathbf{A}\mathbf{x} + \mathbf{b}) - (\mathbf{A}\boldsymbol{\mu} + \mathbf{b}))^\intercal) \\ &= \mathbb{E}([\mathbf{A}(\mathbf{x} - \boldsymbol{\mu})] [\mathbf{A}(\mathbf{x} - \boldsymbol{\mu})]^\intercal) \\ &= \mathbb{E}(\mathbf{A}(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^\intercal \mathbf{A}^\intercal) \\ &= \mathbf{A}\mathbb{E}((\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^\intercal)\mathbf{A}^\intercal \\ &= \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\intercal \end{aligned} \tag{16}$$

1.11 Conditional probability

조건부 확률은 **두 사건 A,B가 있을 때 A가 발생했을 때 B가 발생할 확률을 의미**한다. 이는 다음과 같이 나타낼 수 있다.

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \tag{17}$$

이를 통해 두 사건이 동시에 발생한 확률은 $P(B \cap A) = P(A)P(B|A)$와 같이 나타낼 수 있다. 만약 두 사건 A,B가 독립사건이면 조건부 확률은 다음과 같다.

$$P(B|A) = P(B)P(A) \tag{18}$$

1.12 Bayesian rule

Bayesian Rule은 다음과 같이 조건부확률 간 관계를 의미한다.

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} \tag{19}$$

이 때 $P(B|A)$를 Posterior라고 하고 $P(A|B)$를 Likelihood, $P(B)$를 Prior라고 한다. 예를 들어, 로봇의 위치를 $\mathbf{x}$, 로봇의 센서를 통해 관측한 값을 $\mathbf{z}$이라고 했을 때 주어진 관측 데이터를 바탕으로 현재 로봇이 $\mathbf{x}$에 위치할 확률 $p(\mathbf{x}|\mathbf{z})$는

$$p(\mathbf{x}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{z})} = \eta \cdot p(\mathbf{z}|\mathbf{x})p(\mathbf{x}) \tag{20}$$

과 같다. $p(\mathbf{z}|\mathbf{x})$는 $\mathbf{x}$ 위치에서 해당 관측값이 나올 확률을 의미하며 마지막으로 $p(\mathbf{x})$는 로봇이 $\mathbf{x}$ 위치에 존재할 확률을 의미한다. $1/p(\mathbf{z})$는 η로 치환하여 주로 표기하며 $p(\mathbf{z}|\mathbf{x})p(\mathbf{x})$ 확률분포의 넓이가 1이 되도록 정규화해주는 normalization factor라고 한다.

1.13 Conditional gaussian distribution

두 개의 확률변수 a, b가 주어졌을 때 조건부 확률분포 $p(a|b = b_0)$가 가우시안 분포를 따른다고 하면

$$\begin{aligned} p(a|b = b_0) &= \frac{p(a, b)}{p(b)} = \frac{p(b|a)p(a)}{p(b)} = \eta \cdot p(a, b)p(a) \\ &\sim \mathcal{N}(\boldsymbol{\mu}_{a|b}, \boldsymbol{\Sigma}_{a|b}) \end{aligned} \tag{21}$$

가 된다 이 때 평균 $\boldsymbol{\mu}_{a|b}$과 분산 $\boldsymbol{\Sigma}_{a|b}$은 아래와 같다.

$$\begin{aligned} \boldsymbol{\mu}_{a|b} &= \boldsymbol{\mu}_a + \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}(b_0 - \boldsymbol{\mu}_b) \\ \boldsymbol{\Sigma}_{a|b} &= \boldsymbol{\Sigma}_{aa} - \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}\boldsymbol{\Sigma}_{ab}^{\intercal} \end{aligned} \tag{22}$$

1.14 Recursive bayes filter

시간 t에 대하여 로봇의 위치에 대한 상태 $\mathbf{x}_t$와 관측값 $\mathbf{z}_t$ 그리고 제어입력 $\mathbf{u}_t$가 주어졌을 때 $\text{bel}(\mathbf{x}_t)$(Belief of $\mathbf{x}_t$)는 다음과 같이 정의한다.

$$\text{bel}(\mathbf{x}_t) = p(\mathbf{x}_t \mid \mathbf{z}_{1:t}, \mathbf{u}_{1:t}) \tag{23}$$

$\text{bel}(\mathbf{x}_t)$는 처음부터 t초까지 센서를 통한 관측과 제어입력으로 인해 현재 로봇이 $\mathbf{x}$에 위치할 확률을 의미한다. 해당 식을 Markov Assumption과 Bayesian Rule을 사용하여 전개하면 재귀적인 필터를 유도할 수 있고 이를 Recursive Bayes Filter라고 한다. Recursive Bayes Filter 다음과 같다.

$$\begin{aligned} \text{bel}(\mathbf{x}_t) &= \eta \cdot p(\mathbf{z}_t \mid \mathbf{x}_t)\overline{\text{bel}}(\mathbf{x}_t) \\ \overline{\text{bel}}(\mathbf{x}_t) &= \int p(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \mathbf{u}_t)\text{bel}(\mathbf{x}_{t-1}) \end{aligned} \tag{24}$$

Recursive Bayes Filter는 위와 같이 이전 스텝의 $\text{bel}(\mathbf{x}_{t-1})$로부터 현재 스텝의 $\text{bel}(\mathbf{x}_t)$를 구할 수 있으므로 재귀 필터라고 부른다. **이 때, $p(\mathbf{z}_t \mid \mathbf{x}_t)$를 관측 모델(observation model)이라고 부르며 $\int p(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \mathbf{u}_t)$를 모션 모델(motion model)이라고 부른다. $\text{bel}(\mathbf{x}_t)$가 가우시안 분포를 따르는 경우 이를 특별히 칼만 필터(kalman filter)라고 한다.**

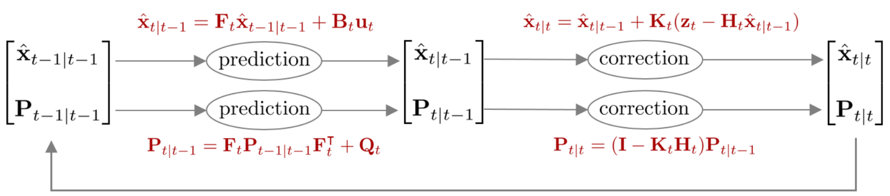
$$\overline{\text{bel}}(\mathbf{x}_t) \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_{t|t-1}, \boldsymbol{\Sigma}_{t|t-1}) \quad \text{(Kalman Filter Prediction)}$$
$$\text{bel}(\mathbf{x}_t) \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_{t|t}, \boldsymbol{\Sigma}_{t|t}) \quad \text{(Kalman Filter Correction)} \tag{25}$$

평균과 분산은 $(\hat{\boldsymbol{\mu}}, \boldsymbol{\Sigma})$ 또는 $(\hat{\mathbf{x}}, \mathbf{P})$로 표현하기도 한다. 이는 표기법만 다를 뿐 동일한 의미를 지닌다. 다음 섹션에서는 $(\hat{\mathbf{x}}, \mathbf{P})$ 표기법을 사용하여 평균과 분산을 나타낸다.

2 Kalman filter (KF)

NOMENCLATURE of kalman filter

- 스칼라는 일반 소문자로 표기한다 e.g., a
- 벡터는 굵은 소문자로 표기한다 e.g., $\mathbf{x}$
- 행렬은 굵은(bold) 대문자로 표기한다 e.g., $\mathbf{F}$
- prediction: $\overline{\text{bel}}(\mathbf{x}_t)$의 평균과 공분산은 $(\hat{\mathbf{x}}_{t|t-1}, \mathbf{P}_{t|t-1})$로 표기한다
  - mean and cov at t given $t-1$
  - 이는 $t-1$ 스텝까지 값이 주어졌을 때 t 스텝에서 평균과 분산을 의미한다.
- correction: $\text{bel}(\mathbf{x}_t)$의 평균과 공분산은 $(\hat{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t})$로 표기한다.
  - mean and cov at t given t
  - 이는 t 스텝까지 값이 주어졌을 때 t 스텝에서 평균과 분산을 의미한다.
- $\mathbf{x}_t$: 모델의 상태 변수(state variable)
- $\mathbf{u}_t$: 모델의 입력(input)
- $\mathbf{F}_t$: 모델의 상태천이(state transition) 행렬
- $\mathbf{B}_t$: 모델에 입력 $\mathbf{u}_t$가 주어졌을 때 $\mathbf{u}_t$를 상태 변수로 변환해주는 행렬
- $\mathbf{H}_t$: 모델의 관측(observation) 행렬
- $\mathbf{w}_t \sim \mathcal{N}(0, \mathbf{Q}_t)$: 모션 모델의 노이즈. $\mathbf{Q}_t$는 $\mathbf{w}_t$의 공분산 행렬을 의미한다.
- $\mathbf{v}_t \sim \mathcal{N}(0, \mathbf{R}_t)$: 관측 모델의 노이즈. $\mathbf{R}_t$는 $\mathbf{v}_t$의 공분산 행렬을 의미한다.



시간 t에 로봇의 위치를 $\mathbf{x}_t$, 로봇의 센서로 부터 관측한 값을 $\mathbf{z}_t$, 로봇의 제어입력을 $\mathbf{u}_t$라고 하면 이를 통해 모션 모델(motion model)과 관측 모델(observation model)을 정의할 수 있다. **이 때, 모션 모델과 관측 모델은 선형이어야(linear model) 한다는 제약조건이 있다.** 모션 모델과 관측 모델은 다음과 같다.

$$\begin{aligned} &\text{Motion Model:} && \mathbf{x}_t = \mathbf{F}_t\mathbf{x}_{t-1} + \mathbf{B}_t\mathbf{u}_t + \mathbf{w}_t \\ &\text{Observation Model:} && \mathbf{z}_t = \mathbf{H}_t\mathbf{x}_t + \mathbf{v}_t \end{aligned} \tag{26}$$

시간 t에서 확률 변수가 모두 가우시안 분포를 따른다고 가정하면 $p(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \mathbf{u}_t), p(\mathbf{z}_t \mid \mathbf{x}_t)$는 다음과 같이 나타낼 수 있다.

$$\begin{aligned} p(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \mathbf{u}_t) \quad & \sim \mathcal{N}(\mathbf{F}_t\mathbf{x}_{t-1} + \mathbf{B}_t\mathbf{u}_t, \mathbf{Q}_t) \\ & = \frac{1}{\sqrt{\det(2\pi\mathbf{Q}_t)}} \exp\Big(-\frac{1}{2}(\mathbf{x}_t - \mathbf{F}_t\mathbf{x}_{t-1} - \mathbf{B}_t\mathbf{u}_t)^\intercal \mathbf{Q}_t^{-1}(\mathbf{x}_t - \mathbf{F}_t\mathbf{x}_{t-1} - \mathbf{B}_t\mathbf{u}_t)\Big) \\ p(\mathbf{z}_t \mid \mathbf{x}_t) \quad & \sim \mathcal{N}(\mathbf{H}_t\mathbf{x}_t, \mathbf{R}_t) \\ & = \frac{1}{\sqrt{\det(2\pi\mathbf{R}_t)}} \exp\Big(-\frac{1}{2}(\mathbf{z}_t - \mathbf{H}_t\mathbf{x}_t)^\intercal \mathbf{R}_t^{-1}(\mathbf{z}_t - \mathbf{H}_t\mathbf{x}_t)\Big) \end{aligned} \tag{27}$$

$\mathbf{Q}$는 모션 모델의 노이즈를 의미하며 $\mathbf{R}$은 관측 모델의 노이즈를 의미한다. 다음으로 칼만 필터를 통해 구해야 하는 $\overline{\text{bel}}(\mathbf{x}_t), \text{bel}(\mathbf{x}_t)$은 아래와 같이 나타낼 수 있다.

$$\begin{aligned} \overline{\text{bel}}(\mathbf{x}_t) &= \int p(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \mathbf{u}_t)\text{bel}(\mathbf{x}_{t-1}) \sim \mathcal{N}(\hat{\mathbf{x}}_{t|t-1}, \mathbf{P}_{t|t-1}) \\ \text{bel}(\mathbf{x}_t) &= \eta \cdot p(\mathbf{z}_t \mid \mathbf{x}_t)\overline{\text{bel}}(\mathbf{x}_t) \sim \mathcal{N}(\hat{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t}) \end{aligned} \tag{28}$$

칼만 필터는 prediction에서 이전 스텝의 값과 모션 모델을 사용하여 예측값 $\overline{\text{bel}}(\mathbf{x}_t)$을 먼저 구한 후 correction에서 관측값과 관측 모델을 사용하여 보정된 값 $\text{bel}(\mathbf{x}_t)$을 구하는 방식으로 동작한다. 초기값 $\text{bel}(\mathbf{x}_0)$은 다음과 같이 주어진다.

$$\text{bel}(\mathbf{x}_0) \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0) \tag{29}$$

2.1 Prediction step

Prediction은 $\overline{\text{bel}}(\mathbf{x}_t)$를 구하는 과정을 말한다. $\overline{\text{bel}}(\mathbf{x}_t)$는 가우시안 분포를 따르므로 평균 $\hat{\mathbf{x}}_{t|t-1}$과 분산 $\mathbf{P}_{t|t-1}$을 각각 구해보면 아래와 같이 구할 수 있다.

$$\boxed{\begin{aligned} \hat{\mathbf{x}}_{t|t-1} &= \mathbf{F}_t\hat{\mathbf{x}}_{t-1|t-1} + \mathbf{B}_t\mathbf{u}_t \\ \mathbf{P}_{t|t-1} &= \mathbf{F}_t\mathbf{P}_{t-1|t-1}\mathbf{F}_t^\intercal + \mathbf{Q}_t \end{aligned}} \tag{30}$$

2.2 Correction step

Correction은 $\text{bel}(\mathbf{x}_t)$를 구하는 과정을 말한다. $\text{bel}(\mathbf{x}_t)$ 또한 가우시안 분포를 따르므로 평균 $\hat{\mathbf{x}}_{t|t}$과 분산 $\mathbf{P}_{t|t}$을 각각 구해보면 다음과 같다. 이 때, $\mathbf{K}_t$는 칼만 게인(kalman gain)을 의미한다.

$$\boxed{\begin{aligned} \mathbf{K}_t &= \mathbf{P}_{t|t-1}\mathbf{H}_t^\intercal(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^\intercal + \mathbf{R}_t)^{-1} \\ \hat{\mathbf{x}}_{t|t} &= \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1}) \\ \mathbf{P}_{t|t} &= (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1} \end{aligned}} \tag{31}$$

2.3 Summary

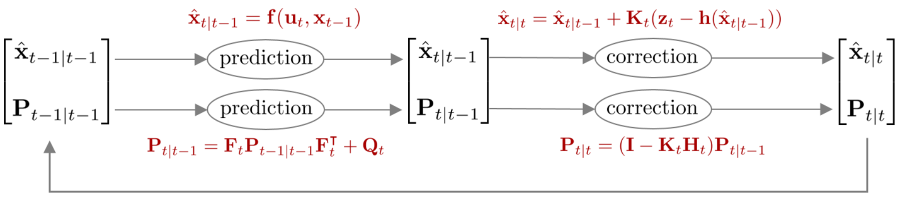
칼만 필터를 함수로 표현하면 다음과 같다.

$$\boxed{\begin{aligned} & \text{KalmanFilter}(\hat{\mathbf{x}}_{t-1|t-1}, \mathbf{P}_{t-1|t-1}, \mathbf{u}_t, \mathbf{z}_t)\{ \\ & \qquad \text{(Prediction Step)} \\ & \qquad \hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t\hat{\mathbf{x}}_{t-1|t-1} + \mathbf{B}_t\mathbf{u}_t \\ & \qquad \mathbf{P}_{t|t-1} = \mathbf{F}_t\mathbf{P}_{t-1|t-1}\mathbf{F}_t^\intercal + \mathbf{Q}_t \\ & \\ & \qquad \text{(Correction Step)} \\ & \qquad \mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^\intercal(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^\intercal + \mathbf{R}_t)^{-1} \\ & \qquad \hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1}) \\ & \qquad \mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1} \\ & \qquad \text{return} \ \ \hat{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t} \\ & \} \end{aligned}} \tag{32}$$

3 Extended kalman filter (EKF)

NOMENCLATURE of extended kalman filter

- $\mathbf{f}(\cdot)$: 상태 변수의 비선형 모션(motion) 모델 함수
- $\mathbf{h}(\cdot)$: 상태 변수의 비선형 관측(observation) 모델 함수



칼만 필터는 모션 모델과 관측 모델이 선형이라는 가정 하에 상태를 추정한다. 하지만 현실세계의 대부분의 현상들은 비선형으로 모델링되므로 앞서 정의한 칼만 필터를 그대로 적용하면 정상적으로 동작하지 않는다. 비선형의 모션 모델과 관측 모델에서도 칼만 필터를 사용하기 위해 확장칼만필터(extended kalman filter, EKF)가 제안되었다. **EKF는 테일러 1차 근사(talyor 1st approximation)을 사용하여 비선형 모델을 선형 모델로 근사한 후 칼만 필터를 적용하는 방법을 사용한다.** EKF의 모션 모델과 관측 모델은 다음과 같다.

$$\begin{aligned} &\text{Motion Model:} && \mathbf{x}_t = \mathbf{f}(\mathbf{u}_t, \mathbf{x}_{t-1}) + \mathbf{w}_t \\ &\text{Observation Model:} && \mathbf{z}_t = \mathbf{h}(\mathbf{x}_t) + \mathbf{v}_t \end{aligned} \tag{33}$$

위 식에서 $\mathbf{f}(\cdot,\cdot)$은 비선형 모션 모델을 의미하고 $\mathbf{h}(\cdot)$은 비선형 관측 모델을 의미한다. $\mathbf{f}(\cdot), \mathbf{h}(\cdot)$에 각각 1차 테일러 근사를 수행하면 다음과 같다.

$$\begin{aligned} &\mathbf{x}_t \approx \mathbf{f}(\mathbf{u}_t, \hat{\mathbf{x}}_{t-1|t-1}) + \mathbf{F}_t(\mathbf{x}_{t-1} - \hat{\mathbf{x}}_{t-1|t-1}) + \mathbf{w}_t \\ &\mathbf{z}_t \approx \mathbf{h}(\hat{\mathbf{x}}_{t|t-1}) + \mathbf{H}_t(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}) + \mathbf{v}_t \end{aligned} \tag{34}$$

이 때, $\mathbf{F}_t$는 모션 모델의 자코비안 행렬을 의미하며 $\mathbf{H}_t$는 관측 모델의 자코비안 행렬을 의미한다.

$$\boxed{\mathbf{F}_t = \frac{\partial \mathbf{f}(\mathbf{u}_t, \hat{\mathbf{x}}_{t-1|t-1})}{\partial \mathbf{x}_{t-1}}\Big|_{\mathbf{x}_{t-1}=\hat{\mathbf{x}}_{t-1|t-1}} \qquad \mathbf{H}_t = \frac{\partial \mathbf{h}(\hat{\mathbf{x}}_{t|t-1})}{\partial \mathbf{x}_t}\Big|_{\mathbf{x}_t=\hat{\mathbf{x}}_{t|t-1}}} \tag{35}$$

시간 t에서 확률 변수가 모두 가우시안 분포를 따른다고 가정하면 $p(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \mathbf{u}_t), p(\mathbf{z}_t \mid \mathbf{x}_t)$는 다음과 같이 나타낼 수 있다. 표현의 편의를 위해 $\mathbf{f}(\cdot), \mathbf{F}_t(\cdot), \mathbf{h}(\cdot), \mathbf{H}_t(\cdot)$로 표현하였다.

$$\begin{aligned} p(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \mathbf{u}_t) \quad & \sim \mathcal{N}(\mathbf{f}(\cdot) + \mathbf{F}_t(\cdot), \mathbf{Q}_t) \\ & = \frac{1}{\sqrt{\det(2\pi\mathbf{Q}_t)}} \exp\bigg(-\frac{1}{2}(\mathbf{x}_t - \mathbf{f}(\cdot) - \mathbf{F}_t(\cdot))^\intercal \mathbf{Q}_t^{-1}(\mathbf{x}_t - \mathbf{f}(\cdot) - \mathbf{F}_t(\cdot)) \bigg) \\ p(\mathbf{z}_t \mid \mathbf{x}_t) \quad & \sim \mathcal{N}(\mathbf{h}(\cdot) + \mathbf{H}_t(\cdot), \mathbf{R}_t) \\ & = \frac{1}{\sqrt{\det(2\pi\mathbf{R}_t)}} \exp\bigg(-\frac{1}{2}(\mathbf{z}_t - \mathbf{h}(\cdot) - \mathbf{H}_t(\cdot))^\intercal \mathbf{R}_t^{-1}(\mathbf{z}_t - \mathbf{h}(\cdot) - \mathbf{H}_t(\cdot)) \bigg) \end{aligned} \tag{36}$$

$\mathbf{Q}$는 모션 모델의 노이즈를 의미하며 $\mathbf{R}$은 관측 모델의 노이즈를 의미한다. 다음으로 칼만 필터를 통해 구해야 하는 $\overline{\text{bel}}(\mathbf{x}_t), \text{bel}(\mathbf{x}_t)$은 아래와 같이 나타낼 수 있다.

$$\begin{aligned} &\overline{\text{bel}}(\mathbf{x}_t) = \int p(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \mathbf{u}_t)\text{bel}(\mathbf{x}_{t-1}) \sim \mathcal{N}(\hat{\mathbf{x}}_{t|t-1}, \mathbf{P}_{t|t-1}) \\ &\text{bel}(\mathbf{x}_t) = \eta \cdot p(\mathbf{z}_t \mid \mathbf{x}_t)\overline{\text{bel}}(\mathbf{x}_t) \sim \mathcal{N}(\hat{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t}) \end{aligned} \tag{37}$$

EKF 또한 KF와 동일하게 prediction에서 이전 스텝의 값과 모션 모델을 사용하여 예측값 $\overline{\text{bel}}(\mathbf{x}_t)$을 먼저 구한 후 correction에서 관측값과 관측 모델을 사용하여 보정된 값 $\text{bel}(\mathbf{x}_t)$을 구하는 방식으로 동작한다. 초기값 $\text{bel}(\mathbf{x}_0)$은 다음과 같이 주어진다.

$$\text{bel}(\mathbf{x}_0) \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0) \tag{38}$$

3.1 Prediction step

Prediction은 $\overline{\text{bel}}(\mathbf{x}_t)$를 구하는 과정을 말한다. 공분산 행렬을 구할 때 선형화된 자코비안 행렬 $\mathbf{F}_t$가 사용된다.

$$\boxed{\begin{aligned}\hat{\mathbf{x}}_{t|t-1} &= \mathbf{f}(\mathbf{u}_t, \hat{\mathbf{x}}_{t-1|t-1}) \\ \mathbf{P}_{t|t-1} &= \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\intercal + \mathbf{Q}_t\end{aligned}} \tag{39}$$

3.2 Correction step

Correction은 $\text{bel}(\mathbf{x}_t)$를 구하는 과정을 말한다. 칼만 게인과 공분산 행렬을 구할 때 선형화된 자코비안 행렬 $\mathbf{H}_t$가 사용된다.

$$\boxed{\begin{aligned}\mathbf{K}_t &= \mathbf{P}_{t|t-1}\mathbf{H}_t^\intercal(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^\intercal + \mathbf{R}_t)^{-1} \\ \hat{\mathbf{x}}_{t|t} &= \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{h}(\hat{\mathbf{x}}_{t|t-1})) \\ \mathbf{P}_{t|t} &= (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1}\end{aligned}} \tag{40}$$

3.3 Summary

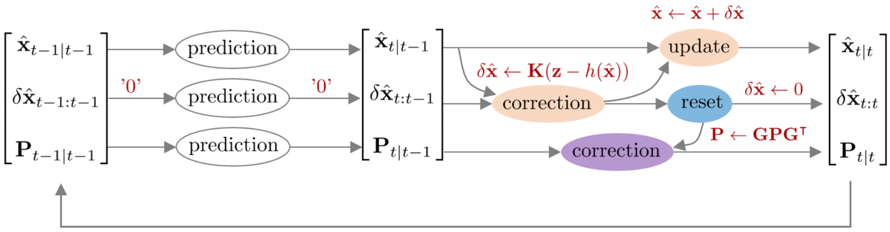
확장칼만필터를 함수로 표현하면 다음과 같다.

$$\boxed{\begin{aligned}&\text{ExtendedKalmanFilter}(\hat{\mathbf{x}}_{t-1|t-1}, \mathbf{P}_{t-1|t-1}, \mathbf{u}_t, \mathbf{z}_t)\{ \\ &\quad \text{(Prediction Step)} \\ &\quad \hat{\mathbf{x}}_{t|t-1} = \mathbf{f}_t(\mathbf{u}_t, \hat{\mathbf{x}}_{t-1|t-1}) \\ &\quad \mathbf{P}_{t|t-1} = \mathbf{F}_t\mathbf{P}_{t-1|t-1}\mathbf{F}_t^\intercal + \mathbf{Q}_t \\ & \\ &\quad \text{(Correction Step)} \\ &\quad \mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^\intercal(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^\intercal + \mathbf{R}_t)^{-1} \\ &\quad \hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{h}_t(\hat{\mathbf{x}}_{t|t-1})) \\ &\quad \mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1} \\ &\quad \text{return} \;\; \hat{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t} \\ &\}\end{aligned}} \tag{41}$$

4 Error-state kalman filter (ESKF)

NOMENCLATURE of error-state kalman filter

- $\mathbf{x}_t$: t 스텝의 상태 변수. true 상태 변수라고도 한다.
- $\hat{\mathbf{x}}_t$: t 스텝의 nominal 상태 변수
- $\delta\mathbf{x}_t$: t 스텝의 에러 상태 변수
- prediction: $\overline{\text{bel}}(\delta\mathbf{x}_t)$의 평균과 공분산은 $(\delta\hat{\mathbf{x}}_{t|t-1}, \mathbf{P}_{t|t-1})$로 표기한다
  - mean and cov at t given $t-1$
  - 이는 $t-1$ 스텝까지 값이 주어졌을 때 t 스텝에서 평균과 분산을 의미한다.
- correction: $\text{bel}(\delta\mathbf{x}_t)$의 평균과 공분산은 $(\delta\hat{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t})$로 표기한다.
  - mean and cov at t given t
  - 이는 t 스텝까지 값이 주어졌을 때 t 스텝에서 평균과 분산을 의미한다.
- $\boldsymbol{\tau}_t \sim \mathcal{N}(0, \mathbf{E}_t)$: 에러 상태 모델의 노이즈. $\mathbf{E}_t$는 $\boldsymbol{\tau}_t$의 공분산 행렬을 의미한다.
- $\mathbf{f}_e(\cdot)$: 에러 상태의 모션 모델 함수. 일반적으로 에러 상태는 작은 값을 가지므로 $\mathbf{f}_e(\cdot)$는 선형성을 가진다.
- $\mathbf{g}(\cdot)$: reset 함수



에러상태 칼만필터(error-state kalman filter, ESKF)는 기존의 상태 변수 $\mathbf{x}_t$의 평균과 분산을 추정하는 EKF와 달리 에러 상태 변수 $\delta\mathbf{x}_t$의 평균과 분산을 추정하는 칼만 필터 알고리즘을 말한다. 이는 error-state extended kalman filter(ES-EKF)라고도 불린다. ESKF는 기존 상태 변수를 true 상태 변수라고 부르며 이를 다음과 같이 nominal 상태와 에러(error) 상태의 합으로 표현한다.

$$\mathbf{x}_t = \hat{\mathbf{x}}_t + \delta\mathbf{x}_t \tag{42}$$

위 식을 해석하면 실제(true) 추정하고자 하는 상태 변수 $\mathbf{x}_t$는 에러가 없는 일반(또는 명목, nominal) 상태 $\hat{\mathbf{x}}_t$과 모델 및 센서 노이즈로 부터 발생하는 에러 상태 $\delta\mathbf{x}_t$의 합으로 나타낼 수 있다는 의미이다. 이 때, nominal 상태는 (상대적으로) 큰 값을 가지며 비선형성을 가진다. 반면에 에러 상태는 0 근처의 작은 값을 가지고 선형성을 가진다. **기존의 EKF는 비선형성이 큰 true (nominal + error) 상태 변수를 선형화하여 필터링하기 때문에 속도가 느리고 시간이 지날수록 에러 값이 누적되는 반면에, ESKF는 에러 상태만을 선형화하여 필터링하기 때문에 속도 및 정확성이 더욱 빠른 장점이 있다.** 기존 EKF와 비교했을 때 ESKF이 가지는 장점들을 정리하면 다음과 같다(Madyastha el tal., 2011):

- 방향(orientation)에 대한 에러 상태 표현법이 최소한의 파라미터를 가진다. 즉, 자유도만큼의 최소 파라미터를 가지기 때문에 over-parameterized로 인해 발생하는 특이점(signularity) 같은 현상이 발생하지 않는다.
- 에러 상태 시스템은 항상 원점(origin) 근처에서만 동작하기 때문에 선형화하기 용이하다. 따라서 짐벌락 같은 파라미터 특이점 현상이 발생하지 않으며 항상 선형화를 수행할 수 있다.
- 에러 상태는 일반적으로 값이 작기 때문에 2차항 이상의 값들은 무시할 수 있다. 이는 자코비안 연산을 쉽고 빠르게 수행할 수 있도록 도와준다. 몇몇 자코비안은 상수화하여 사용하기도 한다.

다만 ESKF는 prediction 속도는 빠르지만 nominal 상태에서 일반적으로 비선형을 가진 큰 값들이 처리되므로 nominal 상태가 처리되는 correction 스텝의 속도가 느린 편이다.

ESKF의 모션 모델과 관측 모델은 다음과 같다.

$$\begin{aligned}
&\text{Motion Model:} && \mathbf{x}_t = \mathbf{f}(\mathbf{x}_{t-1}, \mathbf{u}_t) + \mathbf{w}_t \\
&\text{Error-state Model:} && \delta\mathbf{x}_t = \mathbf{f}_e(\delta\mathbf{x}_{t-1}, \mathbf{u}_t) + \boldsymbol{\tau}_t \\
&\text{Observation Model:} && \mathbf{z}_t = \mathbf{h}(\mathbf{x}_t) + \mathbf{v}_t
\end{aligned} \tag{43}$$

$\mathbf{f}(\cdot)$는 true 상태 방정식의 모션 모델을 의미하고 $\mathbf{f}_e(\cdot)$는 에러 상태 방정식의 모션 모델을 의미한다. $\mathbf{f}_e(\cdot), \mathbf{h}(\cdot)$에 각각 1차 테일러 근사를 수행하면 다음과 같다.

$$\begin{aligned}
\delta\mathbf{x}_t &\approx \mathbf{f}_e(\delta\hat{\mathbf{x}}_{t-1|t-1}, \mathbf{u}_t) + \mathbf{F}_t(\delta\mathbf{x}_{t-1} - \delta\hat{\mathbf{x}}_{t-1|t-1}) + \boldsymbol{\tau}_t \\
\mathbf{z}_t &\approx \mathbf{h}(\hat{\mathbf{x}}_{t|t-1}) + \mathbf{H}_t(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}) + \mathbf{v}_t
\end{aligned} \tag{44}$$

이 때, **$\mathbf{F}_t$는 에러 상태 모션 모델 $\mathbf{f}_e(\cdot)$의 자코비안 행렬을 의미하며 $\mathbf{H}_t$는 에러 상태 관측 모델의 자코비안 행렬을 의미한다. 두 자코비안 모두 true 상태 $\mathbf{x}_t$가 아닌 에러 상태 $\delta\mathbf{x}_t$에 대한 자코비안임에 유의한다. 해당 자코비안 부분이 EKF와 가장 다른 부분이다**

$$\boxed{\mathbf{F}_t = \frac{\partial \mathbf{f}_e(\mathbf{u}_t, \delta\hat{\mathbf{x}}_{t-1|t-1})}{\partial \delta\mathbf{x}_{t-1}}\bigg|_{\delta\mathbf{x}_{t-1}=\delta\hat{\mathbf{x}}_{t-1|t-1}} \qquad \mathbf{H}_t = \frac{\partial \mathbf{h}(\hat{\mathbf{x}}_{t|t-1})}{\partial \delta\mathbf{x}_t}\bigg|_{\delta\mathbf{x}_t=\delta\hat{\mathbf{x}}_{t|t-1}}} \tag{45}$$

$\mathbf{H}_t$는 다음과 같이 연쇄법칙을 통해 표현할 수 있다.

$$\mathbf{H}_t = \frac{\partial \mathbf{h}(\hat{\mathbf{x}}_{t|t-1})}{\partial \delta\mathbf{x}_t} = \frac{\partial \mathbf{h}(\hat{\mathbf{x}}_{t|t-1})}{\partial \mathbf{x}_t}\frac{\partial \mathbf{x}_t}{\partial \delta\mathbf{x}_t} \tag{46}$$

이 중 앞 부분 $\frac{\partial \mathbf{h}(\hat{\mathbf{x}}_{t|t-1})}{\partial \mathbf{x}_t}$은 EKF에서 구한 자코비안과 동일하지만 에러 상태 변수에 대한 자코비안 $\frac{\partial \mathbf{x}_t}{\partial \delta \mathbf{x}_t}$이 추가되었다. 이에 대한 자세한 내용은 Quaternion kinematics for the error-state Kalman filter 내용 정리 포스트를 참고하면 된다.

시간 t에서 확률 변수가 모두 가우시안 분포를 따른다고 가정하면 $p(\delta \mathbf{x}_t \mid \delta \mathbf{x}_{t-1}, \mathbf{u}_t), p(\mathbf{z}_t \mid \delta \mathbf{x}_t)$는 다음과 같이 나타낼 수 있다. 표현의 편의를 위해 $\mathbf{f}_e(\cdot), \mathbf{F}_t(\cdot), \mathbf{h}(\cdot), \mathbf{H}_t(\cdot)$로 표현하였다.

$$
\begin{aligned}
p(\delta \mathbf{x}_t \mid \delta \mathbf{x}_{t-1}, \mathbf{u}_t) \quad & \sim \mathcal{N}(\mathbf{f}_e(\cdot) + \mathbf{F}_t(\cdot), \mathbf{E}_t) \\
& = \frac{1}{\sqrt{\det(2\pi \mathbf{E}_t)}} \exp\bigg(-\frac{1}{2}(\delta \mathbf{x}_t - \mathbf{f}_e(\cdot) - \mathbf{F}_t(\cdot))^\intercal \mathbf{E}_t^{-1} (\delta \mathbf{x}_t - \mathbf{f}_e(\cdot) - \mathbf{F}_t(\cdot)) \bigg) \\
p(\mathbf{z}_t \mid \delta \mathbf{x}_t) \quad & \sim \mathcal{N}(\mathbf{h}(\cdot) + \mathbf{H}_t(\cdot), \mathbf{R}_t) \\
& = \frac{1}{\sqrt{\det(2\pi \mathbf{R}_t)}} \exp\bigg(-\frac{1}{2}(\mathbf{z}_t - \mathbf{h}(\cdot) - \mathbf{H}_t(\cdot))^\intercal \mathbf{R}_t^{-1} (\mathbf{z}_t - \mathbf{h}(\cdot) - \mathbf{H}_t(\cdot)) \bigg)
\end{aligned} \tag{47}
$$

$\mathbf{E}$는 에러 상태 모션 모델의 노이즈를 의미하며 $\mathbf{R}$은 관측 모델의 노이즈를 의미한다. 다음으로 칼만 필터를 통해 구해야 하는 $\overline{\text{bel}}(\delta \mathbf{x}_t), \text{bel}(\delta \mathbf{x}_t)$은 아래와 같이 나타낼 수 있다.

$$
\begin{aligned}
\overline{\text{bel}}(\delta \mathbf{x}_t) &= \int p(\delta \mathbf{x}_t \mid \delta \mathbf{x}_{t-1}, \mathbf{u}_t) \text{bel}(\delta \mathbf{x}_{t-1}) \sim \mathcal{N}(\delta \hat{\mathbf{x}}_{t|t-1}, \mathbf{P}_{t|t-1}) \\
\text{bel}(\delta \mathbf{x}_t) &= \eta \cdot p(\mathbf{z}_t \mid \delta \mathbf{x}_t) \overline{\text{bel}}(\delta \mathbf{x}_t) \sim \mathcal{N}(\delta \hat{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t})
\end{aligned} \tag{48}
$$

ESKF 또한 EKF와 동일하게 prediction에서 이전 스텝의 값과 모션 모델을 사용하여 예측값 $\overline{\text{bel}}(\delta \mathbf{x}_t)$을 먼저 구한 후 correction에서 관측값과 관측 모델을 사용하여 보정된 값 $\text{bel}(\delta \mathbf{x}_t)$을 구하는 방식으로 동작한다. 초기값 $\text{bel}(\delta \mathbf{x}_0)$은 다음과 같이 주어진다.

$$
\text{bel}(\delta \mathbf{x}_0) \sim \mathcal{N}(0, \mathbf{P}_0) \tag{49}
$$

4.1 Prediction step

Prediction은 $\overline{\text{bel}}(\delta \mathbf{x}_t)$를 구하는 과정을 말한다. 공분산 행렬을 구할 때 선형화된 자코비안 행렬 $\mathbf{F}_t$가 사용된다.

$$
\boxed{
\begin{aligned}
\hat{\mathbf{x}}_{t|t-1} &= \mathbf{f}(\mathbf{u}_t, \hat{\mathbf{x}}_{t-1|t-1}) \\
\mathbf{P}_{t|t-1} &= \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^\intercal + \mathbf{Q}_t
\end{aligned}
} \tag{50}
$$

위 식에서 $\mathbf{F}_t$는 true 상태 모션 모델 $\mathbf{f}$에 대한 자코비안이 아닌 에러 상태의 모션 모델 $\mathbf{f}_e$에 대한 자코비안임에 유의한다. 에러 상태 변수의 예측값 $\delta \hat{\mathbf{x}}_{t|t-1}$는 다음과 같이 업데이트 된다.

$$
\delta \hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \delta \hat{\mathbf{x}}_{t-1|t-1} \tag{51}
$$

하지만 매 correction 스텝이 끝나면 reset 함수에 의해 $\delta \hat{\mathbf{x}}$는 0으로 초기화된다. **여기에 선형 자코비안 $\mathbf{F}_t$를 곱해도 값은 0이 되기 때문에 prediction 스텝에서 $\delta \hat{\mathbf{x}}$ 값은 항상 0이 된다.**

4.2 Correction step

Correction은 $\text{bel}(\delta \mathbf{x}_t)$를 구하는 과정을 말한다. 칼만 게인과 공분산 행렬을 구할 때 선형화된 자코비안 행렬 $\mathbf{H}_t$가 사용된다.

$$
\boxed{
\begin{aligned}
& \mathbf{K}_t = \mathbf{P}_{t|t-1} \mathbf{H}_t^\intercal (\mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^\intercal + \mathbf{R}_t)^{-1} \\
& \hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t (\mathbf{z}_t - \mathbf{h}(\hat{\mathbf{x}}_{t|t-1})) \\
& \mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{P}_{t|t-1} \\
& \text{reset } \delta \hat{\mathbf{x}}_{t|t} \\
& \qquad \delta \hat{\mathbf{x}}_{t|t} \leftarrow 0 \\
& \qquad \mathbf{P}_{t|t} \leftarrow \mathbf{G} \mathbf{P}_{t|t} \mathbf{G}^\intercal
\end{aligned}
} \tag{52}
$$

위 식에서 $\mathbf{H}_t$는 관측 모델에 대한 true 상태 $\mathbf{x}_t$의 자코비안이 아닌 에러 상태 $\delta \mathbf{x}_t$의 자코비안임에 유의한다. $\mathbf{P}_{t|t-1}, \mathbf{K}_t$ 또한 EKF와 기호만 같을 뿐 실제 값은 다름에 유의한다. 위 식에서 $\delta \hat{\mathbf{x}}_{t|t}$는 다음과 같다.

$$
\delta \hat{\mathbf{x}}_{t|t} \leftarrow \mathbf{K}_t (\mathbf{z}_t - \mathbf{h}(\hat{\mathbf{x}}_{t|t-1})) \tag{53}
$$

즉, (52)의 두번째 줄이 에러 상태를 칼만 필터를 통해 업데이트한 후 기존 nominal 상태에 더해주는 과정을 의미한다.

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \delta\hat{\mathbf{x}}_{t|t} \tag{54}$$

전체적인 공식은 EKF와 동일하지만 $\mathbf{F}, \mathbf{H}, \mathbf{P}, \mathbf{K}, \mathbf{R}, \mathbf{Q}$ 행렬이 에러 상태 $\delta\mathbf{x}_t$에 대한 값을 의미한다.

4.2.1 Reset

nominal 상태가 정상적으로 업데이트되면 다음으로 에러 상태의 값을 리셋한다. 리셋을 하는 이유는 새로운 nominal 상태에 대한 새로운 에러(new error)를 표현해야 하기 때문이다. ESKF 업데이트의 최종 단계에서는 이러한 새로운 에러를 기반으로 에러 상태의 공분산 $\mathbf{P}_{t|t}$이 업데이트된다.

error 리셋 함수를 $\mathbf{g}(\cdot)$라고 하면 이는 다음과 같이 나타낼 수 있다.

$$\delta\mathbf{x} \leftarrow \mathbf{g}(\delta\mathbf{x}) = \delta\mathbf{x} - \delta\hat{\mathbf{x}}_{t|t} \tag{55}$$

ESKF의 리셋 과정은 다음과 같다.

$$\begin{aligned} \delta\hat{\mathbf{x}}_{t|t} &\leftarrow 0 \\ \mathbf{P}_{t|t} &\leftarrow \mathbf{G}\mathbf{P}_{t|t}\mathbf{G}^{\intercal} \end{aligned} \tag{56}$$

$\mathbf{G}$는 다음과 같이 정의된 리셋에 대한 자코비안을 의미한다.

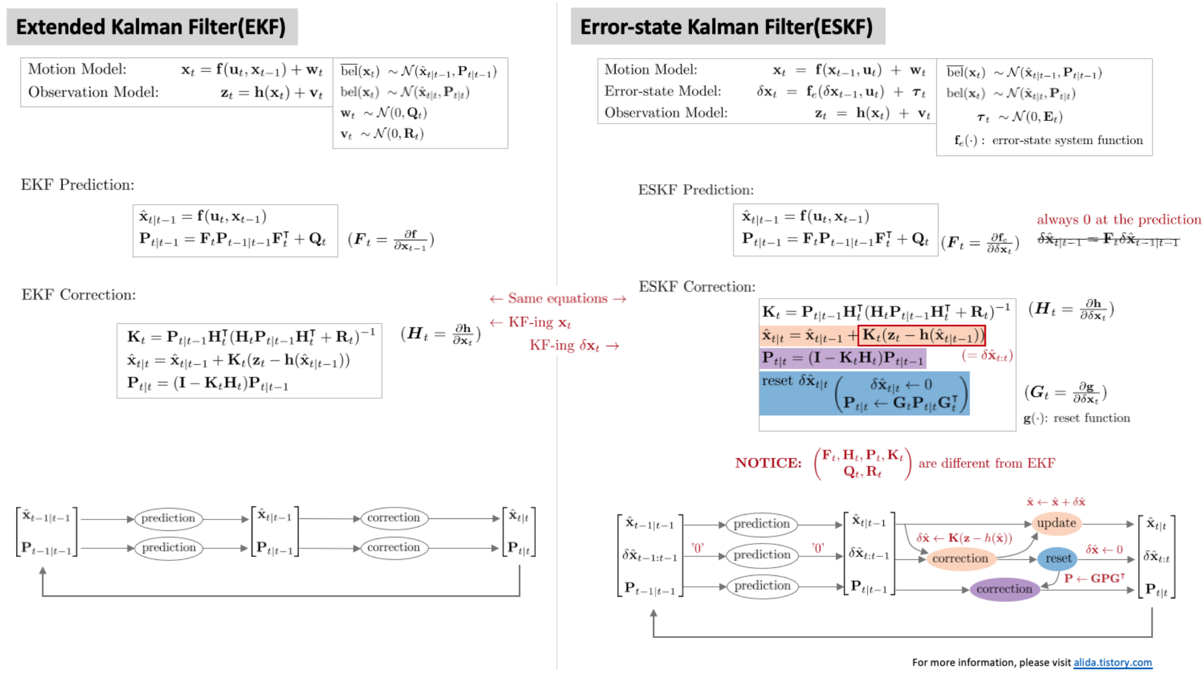
$$\mathbf{G} = \frac{\partial \mathbf{g}}{\partial \delta\mathbf{x}}\bigg|_{\delta\hat{\mathbf{x}}_{t|t}} \tag{57}$$

4.3 Summary

ESKF를 함수로 표현하면 다음과 같다.

$$\boxed{\begin{aligned}
&\text{ErrorStateKalmanFilter}(\delta\hat{\mathbf{x}}_{t-1|t-1},\ \mathbf{P}_{t-1|t-1},\ \mathbf{u}_t,\ \mathbf{z}_t)\ \{ \\
&\qquad \text{(Prediction Step)} \\
&\qquad \hat{\mathbf{x}}_{t|t-1} = \mathbf{f}_t(\mathbf{u}_t, \hat{\mathbf{x}}_{t-1|t-1}) \\
&\qquad \mathbf{P}_{t|t-1} = \mathbf{F}_t\mathbf{P}_{t-1|t-1}\mathbf{F}_t^{\intercal} + \mathbf{Q}_t \\
&\qquad \#\ \delta\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t\hat{\mathbf{x}}_{t-1|t-1} \quad \#\ \text{always } 0 \\
&\\
&\qquad \text{(Correction Step)} \\
&\qquad \mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^{\intercal}(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^{\intercal} + \mathbf{R}_t)^{-1} \\
&\qquad \hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t(\mathbf{z}_t - \mathbf{h}_t(\hat{\mathbf{x}}_{t|t-1})) \\
&\qquad \mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1} \\
&\qquad \text{reset } \delta\hat{\mathbf{x}}_{t|t} \\
&\qquad\quad \delta\hat{\mathbf{x}}_{t|t} \leftarrow 0 \\
&\qquad\quad \mathbf{P}_{t|t} \leftarrow \mathbf{G}\mathbf{P}_{t|t}\mathbf{P}^{\intercal} \\
&\qquad \text{return } \delta\hat{\mathbf{x}}_{t|t},\ \mathbf{P}_{t|t} \\
&\quad \}
\end{aligned}} \tag{58}$$

지금까지 설명한 ESKF와 EKF를 비교하면 아래 그림과 같다.



5 Reference

1. [wiki] Kalman Filter
2. [paper] Sola, Joan. "Quaternion kinematics for the error-state Kalman filter." arXiv preprint arXiv:1711.02508 (2017).
3. [youtube] Robot Mapping Coure - Freiburg Univ
4. [blog] [SLAM] Kalman filter and EKF(Extended Kalman Filter) - Jinyong Jeong

6 Revision log

- 1st: 2020-06-23
- 2nd: 2020-06-24
- 3rd: 2020-06-26
- 4th: 2023-01-21
- 5th: 2023-01-31
- 6th: 2023-02-02
- 7th: 2023-02-04