

Kalman Filter

Mar. 17, 2019 Seo Yeon Yang

Index



- Introduction
- Prediction = Estimation
- Linear Gauss Markov Sequence
- Estimation
- MLP
- Error Update
- Estimation, Update
- Discrete Kalman filter
- Extended Kalman filter
- EKF 의 한계

Introduction



Visual odometry : 순서적인 형태의 짧은 궤적 정보만을 가지고 있음 -> 우리는 전체 운동 궤적이 오랫동안 최적 상태로 유지되기를 바람

Backend optimization

$$\begin{cases} x_k = f(x_{k-1}, u_k) + w_k \\ z_{k,j} = h(y_j, x_k) + v_{k,j} \end{cases} k = 1, \dots, N, j = 1, \dots, M.$$
 (10.1)

Batch process : 장기간 (또는 모든 시간) 상태 추정 문제를 고려하고 과거 정보를 사용하여 상태를 업데이트 할 뿐만 아니라 향후 정보로 업데이트 ->X0, x1, x2, Xn 을 Z0, Z1, Z2,Zn 이 전부 주어질때 추정하는 방식

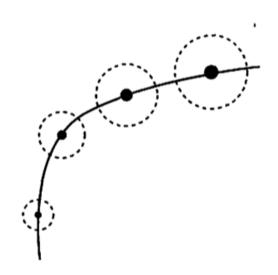
Incremental process : 현재 상태가 과거 순간 또는 바로 이전 순간에 의해서만 결정되는 경우

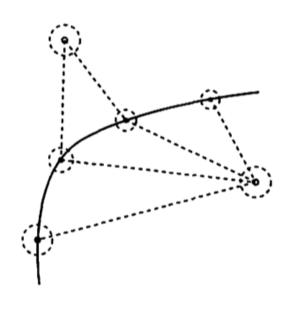
-> Xk 를 X0, X1, X2.... Xk-1 과 Z0, Z1, Z2, ... Zk-1 로 추정하는 방식

♥j는 항상 생성되지 않는다. 움직임을 측정할 장치가 없을 수도 있기 때문에 운동 방정식이 없을 수도 있다. ->몇 가지 관측 방정식으로 구성. SfM (Structure from Motion)

Introduction







운동방정식만을 고려한 경우

관측치를 기반으로 업데이트 해준경우

$$\begin{cases} x_k = f(x_{k-1}, u_k) + w_k \\ z_{k,j} = h(y_j, x_k) + v_{k,j} \end{cases} k = 1, \dots, N, j = 1, \dots, M.$$
 (10.1)

지속적으로 누적되는 error 를 제거하여 update 해준다.

Introduction



운동방정식
$$\begin{cases} x_k = A_k x_{k-1} + u_k + w_k \\ z_k = C_k x_k + v_k \end{cases}$$
 $k = 1, \dots, N.$ (10.9)

Gaussian noise
$$w_k \sim N(\mathbf{0}, \mathbf{R})$$
. $v_k \sim N(\mathbf{0}, \mathbf{Q})$.

(10.10)

Xk 의 상태를 예측하기 위해서

Prediction

$$P(x_k|x_0, u_{1:k}, z_{1:k-1}) = N(A_k \hat{x}_{k-1} + u_k, A_k \hat{P}_{k-1} A_k^{\mathrm{T}} + R).$$
(10.11)

Error Covariance (between xk, xhat k)

$$\bar{\boldsymbol{x}}_k = \boldsymbol{A}_k \hat{\boldsymbol{x}}_{k-1} + \boldsymbol{u}_k,$$

$$\bar{x}_k = A_k \hat{x}_{k-1} + u_k, \quad \bar{P}_k = A_k \hat{P}_{k-1} A_k^{\mathrm{T}} + R.$$

(10.12)

Markov Sequence



Markov Sequence

A random sequence $x(k), k = 0, 1, \dots, N$ is said to be "Markov", if

$$P\left[x(k+1)/x(k),x(k-1),\cdots,x(0)\right] = P\left[x(k+1)/x(k)\right] \text{ for all } k$$
1st order

$$= P\left[x(k+1)/\frac{x(k),x(k-1)}{2^{\text{nd}} \text{ order}}\right] \text{ for all } k$$

ex)
$$x(k+1) = c(k)x(k) + w(k)$$

where $c(k)$: deterministic, $w(k)$: random

→ We can predict tomorrow using the today's information without knowing all the past information!

Linear Gauss Markov Sequence

Linear Gauss Markov Sequence

```
x(k+1) = f(x(k), x(k-1), \dots, x(k-n+1)): Nonlinear Markov Sequence P(k+1/k, k-1, \dots, 0) = P(k+1/k, k-1, \dots, k-(n-1)) \rightarrow n-th order Markov sequence
```

$$x(k+1) = \Phi(k)x(k) + \Gamma(k)w(k), x: n \times 1 \rightarrow \text{Linear 1'st order Markov seq.}$$

$$\Phi(k), \Gamma(k) : \text{ deterministic}$$

$$x(0) \sim N(\overline{x}_0, X_0)$$

$$w(k) \sim N(\overline{w}(k), W(k))$$

$$E\left[\left\{x(0) - \overline{x}(0)\right\}\left\{w(k) - \overline{w}(k)\right\}^T\right] = 0 \text{ for } k = 0, 1, \cdots$$

$$E\left[\left\{w(i) - \overline{w}(i)\right\}\left\{w(j) - \overline{w}(j)\right\}^T\right] = 0 \text{ for } i \neq j$$

Because x(0) and w(k) are Gaussian R.V. $\Rightarrow x(k)$ is also Gaussian R.V.

Therefore, this system can be completely described by

- 1. mean-value sequence, $\overline{x}(k)$
- 2. covariance matrix sequence, $X(k) = E\left[\left\{x(k) \overline{x}(k)\right\}\left\{x(k) \overline{x}(k)\right\}^{T}\right]$

Linear Gauss Markov Sequence

$$\overline{x}(k+1) = E\left[x(k+1)\right]$$

$$= E\left[\Phi(k)x(k) + \Gamma(k)w(k)\right]$$

$$= \Phi(k)E\left[x(k)\right] + \Gamma(k)E\left[w(k)\right]$$

$$\therefore \overline{x}(k+1) = \Phi(k)\overline{x}(k) + \Gamma(k)\overline{w}(k) \text{ where } \overline{x}(0) \text{ : initial condition}$$

$$X(k+1) = E\left[x'(k+1)x'(k+1)^T\right] = E\left[x(k+1) - \overline{x}(k+1)\right]^T$$

$$E\left[x'(k+1)x'(k+1)^T\right] = E\left[x'(k+1)x'(k+1)^T\right] = E\left[x'(k+1) - \overline{x}(k+1)\right]^T$$

$$E\left[x'(k+1)x'(k+1)^T\right] = E\left[x'(k+1)x'(k+1)^T\right] = E\left[x'(k+1)x'(k+1)^T\right]$$

$$\therefore X(k+1) = \Phi(k)X(k)\Phi(k)^{T} + \Gamma(k)W(k)\Gamma(k)^{T}$$

Linear Gauss Markov Sequence

Linear Gauss Markov Sequence

$$x(k+1) = \Phi(k)x(k) + \Gamma(k)w(k)$$

$$\begin{cases}
\Phi(k), \Gamma(k) : \text{ deterministic} \\
x(0) \sim N(\overline{x}_0, X_0) \\
w(k) \sim N(\overline{w}(k), W(k))
\end{cases}$$

$$E\left[\left\{x(0) - \overline{x}(0)\right\}\left\{w(k) - \overline{w}(k)\right\}^T\right] = 0 \text{ for } k = 0, 1, \cdots$$

$$E\left[\left\{w(i) - \overline{w}(i)\right\}\left\{w(j) - \overline{w}(j)\right\}^T\right] = 0 \text{ for } i \neq j$$

$$\bar{x}(k+1) = \Phi(k)\bar{x}(k) + \Gamma(k)\bar{w}(k)$$

$$X(k+1) = \Phi(k)X(k)\Phi(k)^{T} + \Gamma(k)W(k)\Gamma(k)^{T}$$

$$x(0), X(0): \text{ initial condition}$$

Estimation

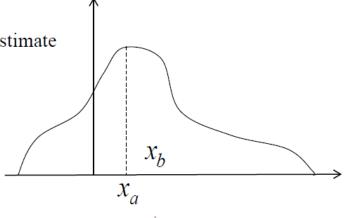


< Criteria for Estimation >

- 1. Maximize probability of x: most probable or likely estimate
- → maximum likelyhood estimator

$$\hat{x} = \text{mode}(= \text{peak}) \text{ of } p(x \mid z) \Rightarrow x_a$$

- 2. Minumum variance: $\int (x \hat{x})^2 p(x \mid z) dx$
- → minimum variance estimator



$$J = \operatorname{var} \left[(x \mid z) \right] = \frac{1}{2} \int_{-\infty}^{\infty} (x - \hat{x})^{2} p(x \mid z) dx \Rightarrow \min_{\hat{x}} J$$

$$\frac{\partial J}{\partial \hat{x}} = 0 : -\int_{-\infty}^{\infty} (x - \hat{x}) p(x \mid z) dx = 0$$

$$\hat{x} \int_{-\infty}^{\infty} p(x \mid z) dx = \int_{-\infty}^{\infty} x p(x \mid z) dx \quad \left(\because \int_{-\infty}^{\infty} p(x \mid z) dx = 1 \right)$$

$$\hat{x} = \int_{-\infty}^{\infty} x p(x \mid z) dx = E\left[(x \mid z) \right] = (x \mid z)_{m} : \text{mean value} \Rightarrow x_{b}$$

Estimation



3. Minimize max. of $|x - \hat{x}|$

→ minimum error estimator

$$\hat{x} = \text{median of } p(x \mid z) \Rightarrow x_c$$

If
$$(x \mid z)$$
: Gaussian $\Rightarrow x_a = x_b = x_c$

∴ Kalman filter (based on Linear and Gaussian distribution) is Min. Variance Min. Error

Max. Likelyhood Min. Variance Estimator!

Estimation



구하고자 하는 것 likelihood

$$P(x_k|x_0, u_{1:k}, z_{1:k-1}) = N(A_k \hat{x}_{k-1} + u_k, A_k \hat{P}_{k-1} A_k^{\mathrm{T}} + R).$$
(10.11)

Posterior

$$P\left(\boldsymbol{z}_{k}|\boldsymbol{x}_{k}\right) = N\left(\boldsymbol{C}_{k}\boldsymbol{x}_{k},\boldsymbol{Q}\right). \tag{10.13}$$

Likelihood can expressed as

$$p(heta|x) = rac{p(x| heta)p(heta)}{p(x)}.^{ extstyle{ beta}}$$

The posterior probability can be written in the memorable form as

Posterior probability \propto Likelihood \times Prior probability.

$$N(\hat{\boldsymbol{x}}_k, \hat{\boldsymbol{P}}_k) = N(\boldsymbol{C}_k \boldsymbol{x}_k, \boldsymbol{Q}) \cdot N(\bar{\boldsymbol{x}}_k, \bar{\boldsymbol{P}}_k). \tag{10.14}$$

$$(x_k - \hat{x}_k)^{\mathrm{T}} \hat{P}_k^{-1} (x_k - \hat{x}_k) = (z_k - C_k x_k)^{\mathrm{T}} Q^{-1} (z_k - C_k x_k) + (x_k - \bar{x}_k)^{\mathrm{T}} \bar{P}_k^{-1} (x_k - \bar{x}_k).$$
(10.15)



Given:
$$x \sim N(\overline{x}, M)$$
 before $z = Hx + v$, $v \sim N(0, V)$ and $E[(x - \overline{x})v^T] = 0$
Find: $x(=\hat{x}) \sim N(\overline{\hat{x}}, P)$ after measurement

$$p(x \mid z) = \frac{p(z \mid x) p(x)}{p(z)}$$

- 1) $x \sim N(\overline{x}, M)$
- 2) $(z \mid x) \sim N(Hx, V)$ (x : must be treated as deterministic)where $\overline{(z \mid x)} = \overline{(Hx + v)} = Hx + \sqrt{v} = Hx$ $E\left[(z - \overline{z})(z - \overline{z})^{T}\right] = E\left[(Hx + v - Hx)(Hx + v - Hx^{T})\right] = E\left[vv^{T}\right] = \sqrt{v}$

3)
$$z \sim N(H\overline{x}, HMH^T + V)$$

where $z = Hx + v$, $\overline{z} = H\overline{x}$ (x: must be treated as stochastic)

$$F\left[\left(z - \overline{z}\right)\left(z - \overline{z}\right)^T\right] = F\left[\left(Hz + v - V - V\right)^T\right] = F\left[\left(Hz + v - V - V\right)^T\right] = F\left[\left(Hz + v - V - V\right)^T\right]$$

$$E\left[\left(z-\frac{1}{z}\right)\left(z-\frac{1}{z}\right)^{T}\right] = E\left[\left(Hx+v-Hz\right)\left(Hx+\nu-Hz\right)^{T}\right] = E\left[H(x-z)b-z^{T}H^{T}+vv^{T}\right]$$



$$p(x) = \frac{1}{(2\pi)^{n/2} |M|^{1/2}} \exp\left[-\frac{1}{2}(x-x)^T M^{-1}(x-x)\right]$$

$$p(x \mid z) = \frac{p(z \mid x) p(x)}{p(z)}$$

$$\max_{x} p(x|z) = \max_{x} \left\{ \exp \begin{bmatrix} -\frac{1}{2} (z - Hx)^{T} V^{-1} (z - Hx) & & & \\ -\frac{1}{2} (x - \overline{x})^{T} M^{-1} (x - \overline{x}) & & & \\ +\frac{1}{2} (z - H\overline{x}) (HMH^{T} + V)^{-1} (z - H\overline{x}) & & & \\ \end{pmatrix} p(z|x) \\ & + \frac{1}{2} (z - H\overline{x}) (HMH^{T} + V)^{-1} (z - H\overline{x}) \right\}$$

$$\Rightarrow \min_{x} J = \frac{1}{2} (z - Hx)^{T} V^{-1} (z - Hx) + \frac{1}{2} (x - \overline{x})^{T} M^{-1} (x - \overline{x})$$
 (1)

$$\rightarrow$$
 Solve x using $\frac{\partial J}{\partial x} = 0$



$$\frac{\partial J}{\partial x} = 0$$

$$(z - Hx)^T V^{-1} (-H) + (x - \overline{x})^T M^{-1} = 0$$
(2)

define
$$(P')^{-1} \triangleq M^{-1} + H^T V^{-1} H$$
 $(P': symmetric)$
$$(x - \overline{x})^T = (z - H \overline{x})^T V^{-1} H P'$$

$$(x - \overline{x}) = P' H^T V^{-1} (z - H \overline{x})$$

$$\therefore \hat{x} = x = \overline{x} + P' H^T V^{-1} (z - H \overline{x})$$
 (4)

 $\delta z = z - H\overline{x}$: residual (If 0, then the previous information is perfect!)



$$\hat{x} = \overline{x} + \left(M^{-1} + H^{T}V^{-1}H\right)^{-1}H^{T}V^{-1}\left(z - H\overline{x}\right)$$
If $M^{-1} = 0$ $(M \to \infty)$ $(: y^{T}M^{-1}y \ge 0$ for any y)
there is no previous information.
$$(2) \to H^{T}V^{-1}(z - Hx) = 0$$

$$(H^{T}V^{-1}H)x = H^{T}V^{-1}z$$

$$\therefore \hat{x} = \left(H^{T}V^{-1}H\right)^{-1}H^{T}V^{-1}z \qquad : \text{Weighted Least Square}$$
If $V = \sigma^{2}I$, $(5) \to \hat{x} = \left(H^{T}H\right)^{-1}H^{T}z$: Least Square

Error Update



```
define e \triangleq \hat{x} - x: estimation error (E(ee^T)?)
e = \overline{x} + P'H^{T}V^{-1}\left(z - H\overline{x}\right) - x
  = \left(\overline{x} - x\right) + P'H^{T}V^{-1}\left(z - H\overline{x}\right) \qquad (z = Hx + v)
  = \left(\overline{x} - x\right) + P'H^{T}V^{-1} \left[v - H\left(\overline{x} - x\right)\right] \qquad (6)
define K \triangleq P'H^TV^{-1}
(6) \rightarrow e = [I - KH](\overline{x} - x) + Kv
P \triangleq E \left[ ee^{T} \right] = E \left[ (I - KH) \left( \overline{x} - x \right) \left( \overline{x} - x \right)^{T} \left( I - KH \right)^{T} + Kvv^{T}K^{T} \right]= (I - KH) M \left( I - KH \right)^{T} + KVK^{T}
  I = P'M^{-1} + P'H^TV^{-1}H (9)
  (9)×M: M = P' + P'H^TV^{-1}HM
                 where K = P'H^TV^{-1}
  P' = (I - KH)M \qquad (10)
```

Error Update



(8):
$$P = P'(I - KH)^T + P'H^TV^{-1}VK^T$$

$$P^{-1} = M^{-1} + H^{T}V^{-1}H$$
 (11)

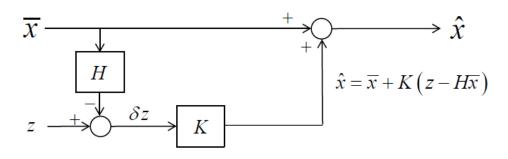
(7): $E[e] = 0 \leftarrow unbiased$

from (3), since $H^TV^{-1}H$ is positive-semidefinite $(y^T(H^TV^{-1}H)y \ge 0 \text{ for any } y)$

$$P^{-1} \ge M^{-1}$$

$$\Rightarrow P \le M \ (i.e. \ y^T P y \le y^T M y \text{ for any } y)$$

 \therefore Uncertainty of x never increase after measurement update!



Discrete Kalman filter



Discrete Kalman Filter

1. Measurement update

$$\hat{x}(k) = \overline{x}(k) + K_d(k) \left[z(k) - H(k) \overline{x}(k) \right] \quad \left(\delta z(k) = z(k) - H(k) \overline{x}(k) \right)$$
where
$$K_d(k) = P(k) H(k)^T V^{-1}(k)$$

$$P(k)^{-1} = M(k)^{-1} + H(k)^T V(k)^{-1} H(k)$$

2. Time update

$$\overline{x}(k+1) = \Phi(k)\hat{x}(k)$$

$$M(k+1) = \Phi(k)P(k)\Phi(k)^{T} + \Gamma_{d}(k)W_{d}(k)\Gamma_{d}(k)^{T}$$

$$\begin{split} x\left(k+1\right) &= \Phi_{d}x\left(k\right) + \Gamma_{d}w_{d}\left(k\right) \\ z\left(k\right) &= Hx\left(k\right) + v\left(k\right) \\ where \ w_{d}\left(k\right) &\sim \mathbf{N}\left(0, W_{d}\right), \ v\left(k\right) \sim \mathbf{N}\left(0, V\right) \\ E\left[\left\{x\left(0\right) - \overline{x}\left(0\right)\right\}w_{d}\left(k\right)^{T}\right] &= 0 \\ E\left[\left\{x\left(k\right) - \overline{x}\left(k\right)\right\}v^{T}\left(k\right)\right] &= 0 \end{split}$$

Estimation, Error Update



$$\begin{cases} \boldsymbol{x}_k = \boldsymbol{A}_k \boldsymbol{x}_{k-1} + \boldsymbol{u}_k + \boldsymbol{w}_k \\ \boldsymbol{z}_k = \boldsymbol{C}_k \boldsymbol{x}_k + \boldsymbol{v}_k \end{cases} \qquad k = 1, \dots, N.$$
 (10.9)

$$\boldsymbol{w}_k \sim N(\boldsymbol{0}, \boldsymbol{R}). \quad \boldsymbol{v}_k \sim N(\boldsymbol{0}, \boldsymbol{Q}).$$
 (10.10)

1. 예측

$$\bar{x}_k = A_k \hat{x}_{k-1} + u_k, \quad \bar{P}_k = A_k \hat{P}_{k-1} A_k^{\mathrm{T}} + R.$$
 (10.24)

2. 업데이트: 칼만 게인이라고도 하는 K를 먼저 계산합니다.

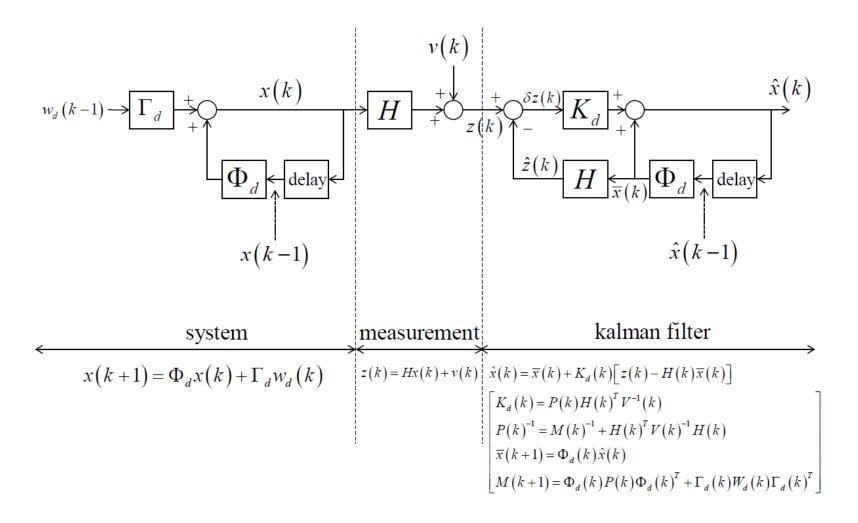
$$\boldsymbol{K} = \bar{\boldsymbol{P}}_{k} \boldsymbol{C}_{k}^{\mathrm{T}} (\boldsymbol{C}_{k} \bar{\boldsymbol{P}}_{k} \boldsymbol{C}_{k}^{\mathrm{T}} + \boldsymbol{Q}_{k})^{-1}. \tag{10.25}$$

그런 다음 사후 확률 분포를 계산하십시오.

$$\hat{\boldsymbol{x}}_{k} = \bar{\boldsymbol{x}}_{k} + \boldsymbol{K} (\boldsymbol{z}_{k} - \boldsymbol{C}_{k} \bar{\boldsymbol{x}}_{k}) \hat{\boldsymbol{P}}_{k} = (\boldsymbol{I} - \boldsymbol{K} \boldsymbol{C}_{k}) \bar{\boldsymbol{P}}_{k}.$$
(10.26)

Discrete Kalman filter





Extended Kalman filter



$$\begin{pmatrix} \dot{x}_N = f(x_N, u_N) + \Gamma w \\ z_N = h(x_N) + v \end{pmatrix}$$

for nominal (or, trim, equilibrium) condition x_{N_0}, u_{N_0}

$$\dot{x}_{N} = \left\{ f\left(x_{N0}, u_{N0}\right) + \left[\frac{\partial f}{\partial x}\Big|_{x_{N0}, u_{N0}}\right] \cdot \delta x + \left[\frac{\partial f}{\partial u}\Big|_{x_{N0}, u_{N0}}\right] \cdot \delta u + h.o.t \right\} + \Gamma w \quad ----- (1)$$

where
$$\dot{x}_{N0} = f\left(x_{N0}, u_{N0}\right)$$

$$\begin{pmatrix} x_N = x_{N0} + \delta x \\ u_N = u_{N0} + \delta u \end{pmatrix}$$

$$\dot{x}_N = \dot{x}_{N0} + \delta \dot{x}$$

from
$$(1),(2)$$
:

$$\delta \dot{x} = F \delta x + G \delta u + \Gamma w \tag{3}$$

where
$$F = \left[\frac{\partial f}{\partial x}\Big|_{x_{N0}, u_{N0}}\right]$$
, $G = \left[\frac{\partial f}{\partial u}\Big|_{x_{N0}, u_{N0}}\right]$

Extended Kalman filter



define
$$x \triangleq \delta x$$
 and $u \triangleq \delta u$ in (3)
$$(3) \Rightarrow \therefore \dot{x} = Fx + Gu + \Gamma w \qquad (4)$$
likewise, $z_N = \left\{ h(x_{N0}) + \left[\frac{\partial h}{\partial x} \Big|_{x_{N0}} \right] \cdot \delta x \right\} + v \qquad (5)$

$$where \begin{cases} z_N = z_{N0} + \delta z \\ z_{N0} = h(x_{N0}) \end{cases}$$

$$\therefore \delta z = Hx + v \qquad \left(where \ H \equiv \left[\frac{\partial h}{\partial x} \Big|_{x_{VO}} \right] \right)$$

define $z \triangleq \delta z$

$$\therefore z = Hx + v$$

EKF 의 한계



• K와 K-1 사이의 관계만을 고려하는 Incremental 방식

-> 현재 상태가 실제로 오래 전의 데이터 (예: 루프 폐쇄)와 관련되어 있으면 필터를 처리하기가 어려울 수 있습니다. 반면에 비선형 최적화 기반 방법은 모든 히스토리 데이터를 사용합니다.

• 선형화로 인한 오차

->EKF 필터는 한 번만 선형화 한 다음 선형화 결과를 기반으로 사후 확률을 직접 계산합니다. 강한 비선형성을 가지면 선형 근사는 매우 작은 범위에서만 유효하며 큰 거리에서 선형성으로 근사화 된 것으로 간주 할 수 없습니다.

Large scale 에 부적합

-> 프로그램 구현 관점에서 볼 때, EKF는 상태 파라미터들 (로봇의 포즈)의 평균과 분산을 저장하고 이를 유지하고 업데이트 해야 합니다. 랜드마크도 상태로 놓으면 시각적 SLAM의 랜드마크가 많기 때문에 저장량은 상당히 크며 상태 파라미터의 수 에 제곱에 해당됩니다.



Thank you



Q&A