Course 04: EKF

Saturday, November 2, 2019 11:02 AM

slam04-ekf

Robot Mapping

Extended Kalman Filter

Cyrill Stachniss

UNI FREIBURG

AIS Autonomous Intelligent Systems

1

goal is to evaluate $p(x|z,u)$ posterior

SLAM is a State Estimation Problem

- Estimate the map and robot's pose
- Bayes filter is one tool for state estimation
- **Prediction**

$$\overline{bel}(x_t) = \int p(x_t \mid u_t, x_{t-1})\, bel(x_{t-1})\, dx_{t-1}$$

- **Correction**

$$bel(x_t) = \eta\, p(z_t \mid x_t)\, \overline{bel}(x_t)$$

2

Kalman Filter

- It is a Bayes filter (as one implementation)
- Estimator for the linear Gaussian case
- Optimal solution for linear models and Gaussian distributions

criteria(1)

criteria(2)

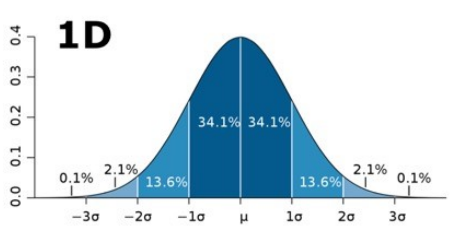
3

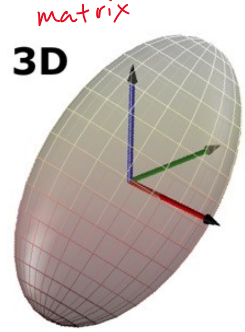
Kalman Filter Distribution

- Everything is Gaussian

$$p(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

covariance matrix





4

?

Properties: Marginalization and Conditioning

x_a, x_b could also be vectors.

- Given $x = \begin{pmatrix} x_a \\ x_b \end{pmatrix}$ $p(x) = \mathcal{N}$ → gaussian

?

- The marginals are Gaussians

$$p(x_a) = \mathcal{N} \qquad p(x_b) = \mathcal{N}$$

- as well as the conditionals

$$p(x_a \mid x_b) = \mathcal{N} \qquad p(x_b \mid x_a) = \mathcal{N}$$

※ (important) if we update our gaussian distributions with sensor model / motion model, the gaussian prop stays.

5

Marginalization

- Given $p(x) = p(x_a, x_b) = \mathcal{N}(\mu, \Sigma)$

$$\text{with} \quad \underset{\text{mean}}{\mu} = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix} \quad \Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$$

- The marginal distribution is

$$p(x_a) = \int p(x_a, x_b)\, dx_b = \mathcal{N}(\underset{\mu_a}{\mu}, \underset{\Sigma_{aa}}{\Sigma})$$

just "cut out" the part, done

$$\text{with} \quad \mu = \mu_a \quad \Sigma = \Sigma_{aa}$$

6

Conditioning

- Given $p(x) = p(x_a, x_b) = \mathcal{N}(\mu, \Sigma)$

$$\text{with} \quad \mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix} \quad \Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$$

- The conditional distribution is

$$p(x_a \mid x_b) = \frac{p(x_a, x_b)}{p(x_b)} = \mathcal{N}(\mu, \Sigma)$$

$$\text{with} \quad \mu = \mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}(b - \mu_b)$$

if this part is large, and I want to know the info of $p(x_a|x_b)$, I HAVE TO INVERT IT, it is a costful manipulati

$$\Sigma = \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}$$

7

if you know very little about
$\Downarrow$
$\Sigma_b \to \infty$
$\Downarrow$
$\Sigma_b^{-1} \to 0$
$\Downarrow$
$p(x_a|x_b) \approx p(x_a$

Linear Model

- The Kalman filter assumes a linear transition and observation model
- Zero mean Gaussian noise

$$x_t = A_t x_{t-1} + B_t u_t + \epsilon_t$$

random noise term

$$z_t = C_t x_t + \delta_t$$

Note 1: here it represents the MEAN of MOTION & OBSERVASION MODELS.
(we'll still have uncertainty associated with that).

8

Components of a Kalman Filter

A_t Matrix $(n \times n)$ that describes how the state evolves from $t-1$ to t without controls or noise.

B_t Matrix $(n \times l)$ that describes how the control u_t changes the state from $t-1$ to t.

C_t Matrix $(k \times n)$ that describes how to map the state x_t to an observation z_t.

ϵ_t δ_t Random variables representing the process and measurement noise that are assumed to be independent and normally distributed with covariance R_t and Q_t respectively.

Note 1: if I have very few knowledge of the world:
MEAN $\to 0$, COVAR $\to \infty$; The more observe the world, the more pec

9

Note 2: $A_t, B_t, C_t, \epsilon_t, \delta_t$ can be recalculated every time t, by collecting more info about the world.

Note 3: in some books R_t & Q_t are flipped; here it stitches to "Probablistic Robotics" Book, $R_t \Leftrightarrow$ control
$Q_t \Leftrightarrow$ observation.

Linear Motion Model

- Motion under Gaussian noise leads to

- Motion under Gaussian noise leads to

$$p(x_t \mid u_t, x_{t-1}) = ?$$

how to estimate where I should be, given u_t & x_{t-1}

$$p(x_t \mid u_t, x_{t-1}) = \underset{\text{something.}}{\underline{\eta}} \exp\left[-\frac{1}{2}\cdot\left(x_t - \left(A_t x_{t-1} + B_t u_t\right)\right) R_t^{-1} (\ldots)\right]$$

10

Linear Motion Model

- Motion under Gaussian noise leads to

$$p(x_t \mid u_t, x_{t-1}) = \det(2\pi R_t)^{-\frac{1}{2}}$$
$$\exp\left(-\frac{1}{2}(x_t - A_t x_{t-1} - B_t u_t)^T R_t^{-1}(x_t - A_t x_{t-1} - B_t u_t)\right)$$

- R_t describes the noise of the motion

11

Linear Observation Model

- Measuring under Gaussian noise leads to

$$p(z_t \mid x_t) = ?$$

$$p(z_t \mid x_t) = \underline{\eta} \exp\left[-\frac{1}{2}\left(z_t - C_t x_t\right) Q_t^{-1} \left(z_t - C_t x_t\right)\right]$$

↳ something.

12

Linear Observation Model

- Measuring under Gaussian noise leads to

$$p(z_t \mid x_t) = \det(2\pi Q_t)^{-\frac{1}{2}}$$
$$\exp\left(-\frac{1}{2}(z_t - C_t x_t)^T Q_t^{-1}(z_t - C_t x_t)\right)$$

- Q_t describes the measurement noise

13

Everything stays Gaussian

- Given an initial Gaussian belief, the belief is always Gaussian

pred
$$\overline{bel}(x_t) = \int \underline{p(x_t \mid u_t, x_{t-1})}\ \underline{bel(x_{t-1})}\, dx_{t-1}$$

corr
$$bel(x_t) = \eta\, \underline{p(z_t \mid x_t)}\ \underline{\overline{bel}(x_t)}$$

⇒ product of two gausians: new MEAN is the product of weighted MEAN.

- Proof is non-trivial TODO: check the proof.
(see Probabilistic Robotics, Sec. 3.2.4)

14

Kalman Filter Algorithm

1: **Kalman_filter**$(\mu_{t-1}, \Sigma_{t-1}, u_t, z_t)$:

predition step

2: $\bar{\mu}_t = A_t\, \mu_{t-1} + B_t\, u_t$

3: $\underbrace{\bar{\Sigma}_t}_{\text{new uncertainty}} = \underbrace{A_t\, \Sigma_{t-1}\, A_t^T}_{\text{processed old uncertainty}} + R_t$

→ we added thru motion uncertainty.

correction step

4: $\underbrace{K_t}_{\text{Kalman gain}} = \bar{\Sigma}_t\, C_t^T (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1}$

Kalman gain : treats how certain am I about the observations . wrt the motion

5: $\mu_t = \bar{\mu}_t + K_t(z_t - C_t\, \bar{\mu}_t)$

6: $\Sigma_t = (I - K_t\, C_t)\, \bar{\Sigma}_t$

7: return μ_t, Σ_t

i.e. motions normally adds uncertain- into the systen

INTUITION :

① what happens if $Q_t = 0 \Rightarrow K_t$

$\mu_t = \bar{\mu}_t + K_t ($

$\| $

C_t^{-1}

$= \bar{\mu}_t + C_t^-$

② - - - - worst

$Q_t \to \infty$

↳ $K_t \to 0$

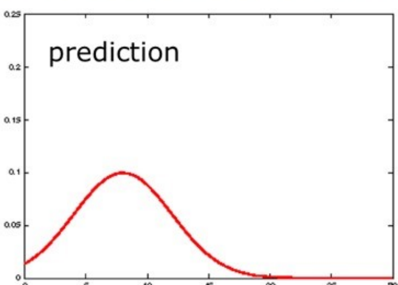
↳ $\mu_t \to \bar{\mu}_t$

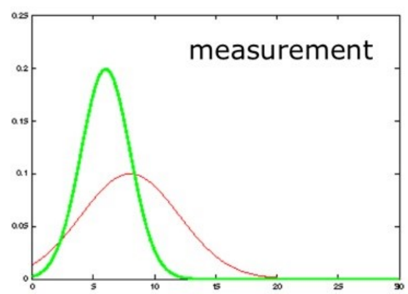
↳ sens

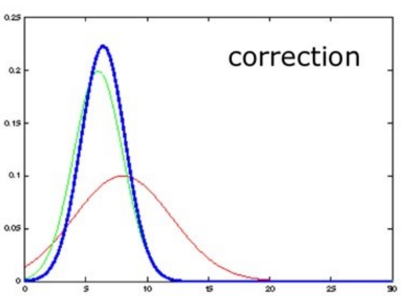
15

※ In a word, the EKF computes weighted mean between prediction & observation

1D Kalman Filter Example (1)



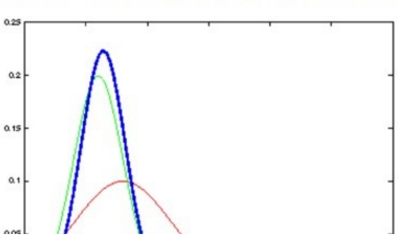


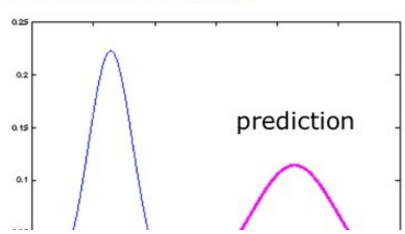


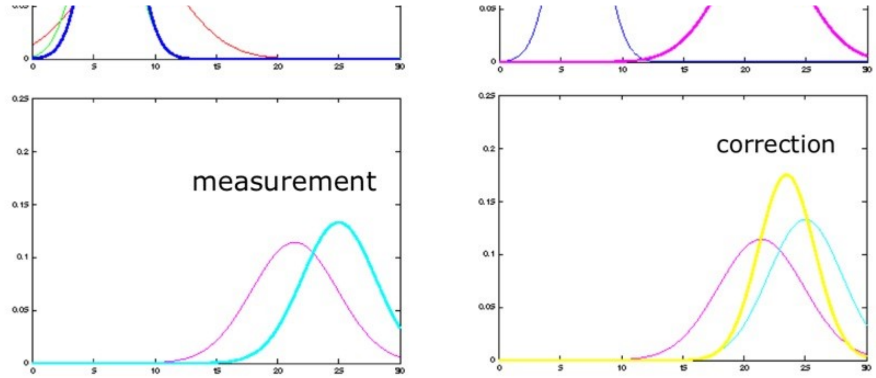
It's a weighted mean!

16 16

1D Kalman Filter Example (2)



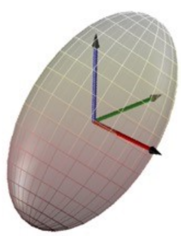




17

Kalman Filter Assumptions

- Gaussian distributions and noise
- Linear motion and observation model



$$x_t = A_t x_{t-1} + B_t u_t + \epsilon_t$$

$$z_t = C_t x_t + \delta_t$$

What if this is not the case?

18

Non-linear Dynamic Systems

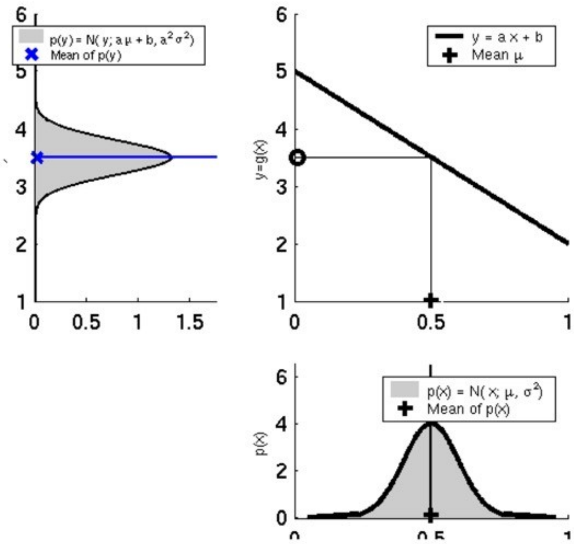
- Most realistic problems (in robotics) involve nonlinear functions

$$\cancel{x_t = A_t x_{t-1} + B_t u_t + \epsilon_t} \qquad \cancel{z_t = C_t x_t + \delta_t}$$

$$x_t = g(u_t, x_{t-1}) + \epsilon_t \qquad z_t = h(x_t) + \delta_t$$

19

Linearity Assumption Revisited



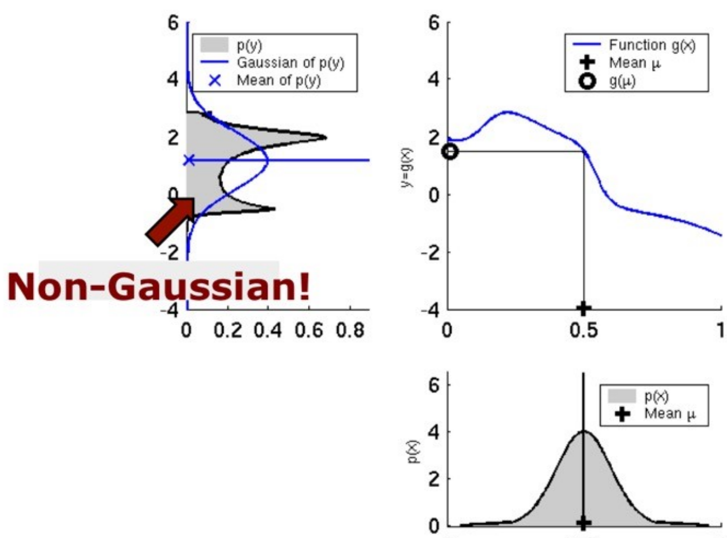
(an important assumption)

⚠ map one state thru linear ~~equ~~ function

↳ the mapped state is also within gaussian dist

20

Non-Linear Function



21

Non-Gaussian Distributions

- The non-linear functions lead to non-

Gaussian distributions

- Kalman filter is not applicable anymore!

What can be done to resolve this?

22

Non-Gaussian Distributions

- The non-linear functions lead to non-Gaussian distributions
- Kalman filter is not applicable anymore!

What can be done to resolve this?

Local linearization!

23

EKF Linearization: First Order Taylor Expansion

- Prediction:

first derivative

$$g(u_t, x_{t-1}) \approx g(u_t, \mu_{t-1}) + \underbrace{\frac{\partial g(u_t, \mu_{t-1})}{\partial x_{t-1}}}_{=: G_t} (x_{t-1} - \mu_{t-1})$$

evaluate the function @ the known best estimate at the moment

how far am I from the best estimate I have. (linearization point).

- Correction:

$$h(x_t) \approx h(\bar{\mu}_t) + \underbrace{\frac{\partial h(\bar{\mu}_t)}{\partial x_t}}_{=:\, H_t} (x_t - \bar{\mu}_t)$$

Jacobian matrices

24

Reminder: Jacobian Matrix

- It is a **non-square matrix** $m \times n$ in general
- Given a vector-valued function

$$g(x) = \begin{pmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x) \end{pmatrix}$$

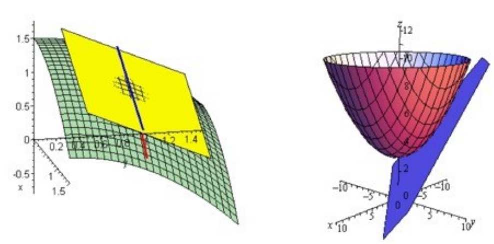
- The **Jacobian matrix** is defined as

$$G_x = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}$$

25

Reminder: Jacobian Matrix

- It is the orientation of the tangent plane to the vector-valued function at a given point



- Generalizes the gradient of a scalar valued function

26

EKF Linearization: First Order Taylor Expansion

- Prediction:

$$g(u_t, x_{t-1}) \approx g(u_t, \mu_{t-1}) \ + \ \underbrace{\frac{\partial g(u_t, \mu_{t-1})}{\partial x_{t-1}}}_{=:\, G_t} (x_{t-1} - \mu_{t-1})$$

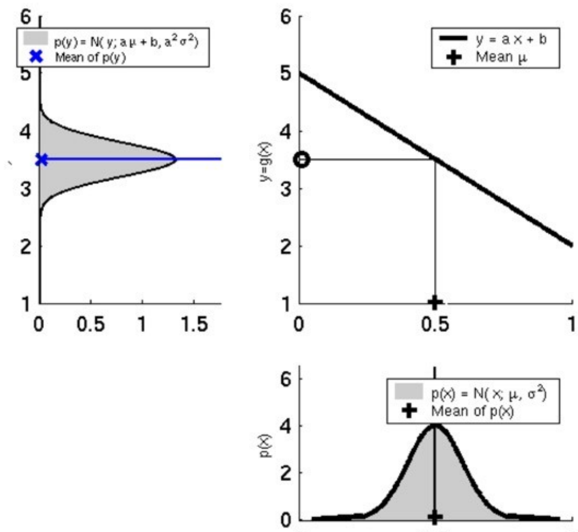
- Correction:

$$h(x_t) \approx h(\bar{\mu}_t) \ + \ \underbrace{\frac{\partial h(\bar{\mu}_t)}{\partial x_t}}_{=:\, H_t} (x_t - \bar{\mu}_t)$$

Linear functions!

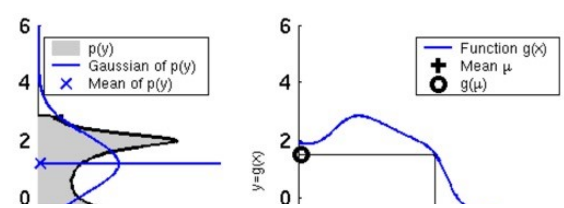
27

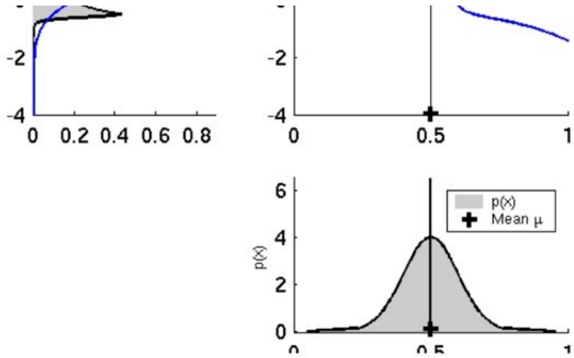
Linearity Assumption Revisited



28

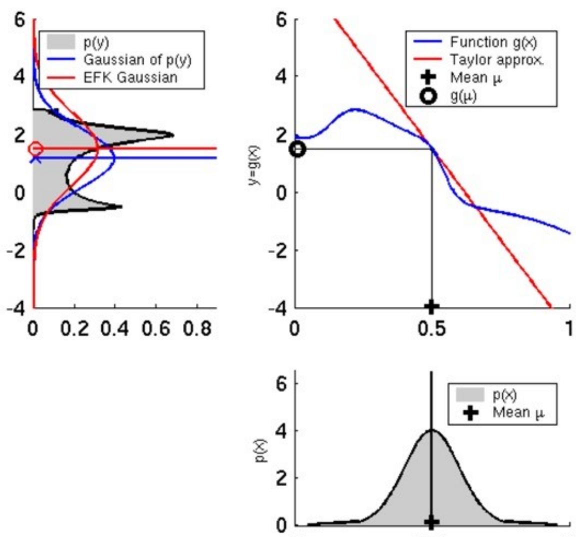
Non-Linear Function





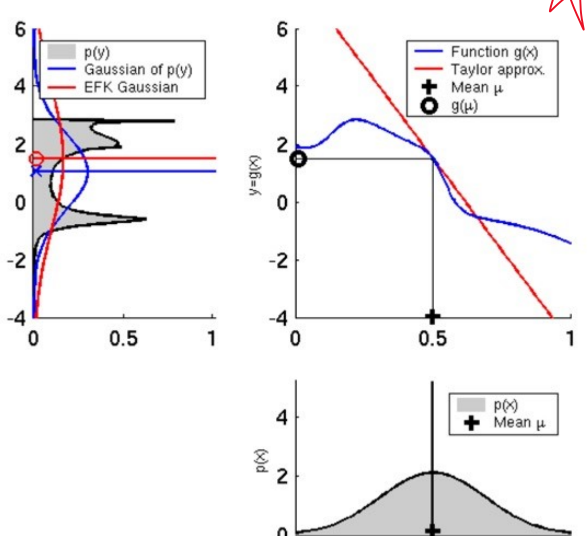
29

EKF Linearization (1)



30

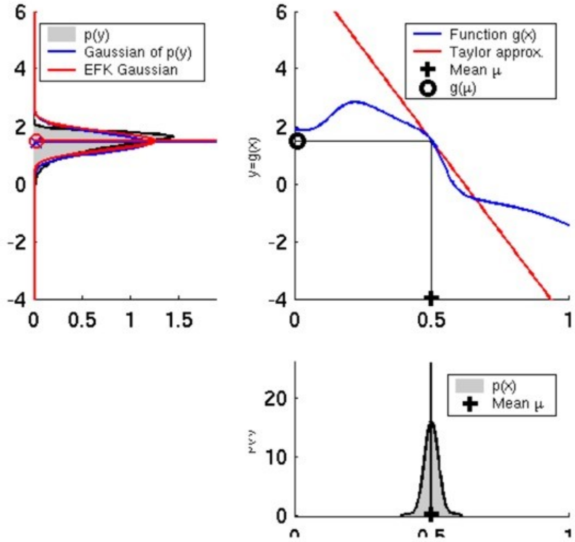
EKF Linearization (2)



Note:
the bigger of the distribution away from MEAN (i.e. linearization point) the worse of the linearization estimation / approxima

31

EKF Linearization (3)



32

Linearized Motion Model

- The linearized model leads to

$$p(x_t \mid u_t, x_{t-1}) \approx \det(2\pi R_t)^{-\frac{1}{2}}$$
$$\exp\Big(-\frac{1}{2}\,(x_t - g(u_t, \mu_{t-1}) - G_t\,(x_{t-1} - \mu_{t-1}))^T$$
$$R_t^{-1}\,(x_t - \underbrace{g(u_t, \mu_{t-1}) - G_t\,(x_{t-1} - \mu_{t-1})}_{\text{linearized model}})\Big)$$

why "$x_{t-1} - \mu_{t-1}$"? are they same?

TODO: check probability book.

- R_t describes the noise of the motion

33

Linearized Observation Model

- The linearized model leads to

$$p(z_t \mid x_t) = \det(2\pi Q_t)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(z_t - h(\bar{\mu}_t) - H_t(x_t - \bar{\mu}_t))^T Q_t^{-1}(z_t - \underbrace{h(\bar{\mu}_t) - H_t(x_t - \bar{\mu}_t)}_{\text{linearized model}})\right)$$

- Q_t describes the measurement noise

34

Extended Kalman Filter Algorithm

1: **Extended_Kalman_filter**($\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$):

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

3: $\bar{\Sigma}_t = G_t \, \Sigma_{t-1} \, G_t^T + R_t$

4: $K_t = \bar{\Sigma}_t \, H_t^T (H_t \, \bar{\Sigma}_t \, H_t^T + Q_t)^{-1}$

5: $\mu_t = \bar{\mu}_t + K_t(z_t - h(\bar{\mu}_t))$

6: $\Sigma_t = (I - K_t \, H_t) \, \bar{\Sigma}_t$

7: return μ_t, Σ_t

$\boxed{A_t \leftrightarrow G_t \quad C_t \leftrightarrow H_t}$

KF vs. EKF

35

Note 1: very similar to KF, only the Gt, Ht changed;
Note 2: every time it needs to be linearized @ new linearization point

Extended Kalman Filter Summary

- Extension of the Kalman filter
- One way to handle the non-linearities
- Performs local linearizations
- Works well in practice for moderate non-linearities

- Large uncertainty leads to increased approximation error error

- Complexity : $O(k^{2.4} + n^2)$ dim of state
  1) 2.4 is due to inversion, fastest, (stupid way is $O(k^3)$)
  2) k is the dim of observation

36

Literature

Kalman Filter and EKF

- Thrun et al.: "Probabilistic Robotics", Chapter 3
- Schön and Lindsten: "Manipulating the Multivariate Gaussian Density"
- Welch and Bishop: "Kalman Filter Tutorial"

37