# L06: EKF and Localization

Gonzalo Ferrer

14 February 2022

## 1 Extended Kalman filter

### Kalman filter: Linear system plus Gaussian prior

$$\left.\begin{aligned} \bar{\mu}_t &= A_t \mu_t + B_t u_t \\ \bar{\Sigma}_t &= A_t \Sigma_{t-1} A_t^T + R_t \end{aligned}\right\} \text{prediction (marginalize)}$$

$$\left.\begin{aligned} K_t &= \bar{\Sigma}_t C_t^T \left(C_t \bar{\Sigma}_t C_t^T + Q\right)^{-1} \\ \mu_t &= \bar{\mu}_t + K_t \left(z_t - C_t \bar{\mu}_t\right) \\ \Sigma_t &= \left(I - K_t C_t\right) \bar{\Sigma}_t \right) \end{aligned}\right\} \text{correction (conditioning)}$$

### Motion model: first order Taylor expansion

$$x_t = g(x_{t-1}, u_t, \epsilon_t) \approx g(\mu_{t-1}, u_t, 0) + \frac{\partial g}{\partial x_{t-1}}\Big|_{\mu_{t-1}} (x_{t-1} - \mu_{t-1}) + \epsilon_t$$

In L05 discussed on how to model $g(\cdot)$ for different systems and how to obtain the probabilistic model.

### Sensor model

We observe features of landmarks (L05):

$$z_t = h(x_t, \eta_t) \approx h(\mu_t, 0) + \frac{\partial h}{\partial x_t}\Big|_{\mu_t} (x_t - \mu_t) + \eta_t$$

### Intuition on linearization

Linearizing assumes **errors**!

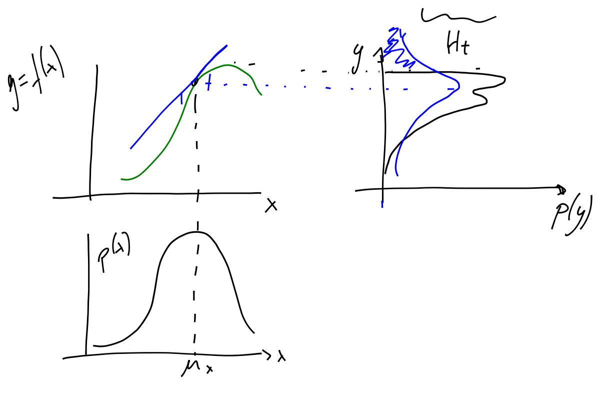


Figure 1: Linearization intuition. The random variable $x \sim p(x)$ is transformed by the function $y = f(x)$. The true distribution of $y \sim p(y)$ would not be Gaussian (black line in the top right), but after linearizing, the PDF is approximated as Gaussian.

### Equations of the Extended Kalman Filter

Inputs: $\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$

1. $\bar{\mu}_t = g(\mu_{t-1}, u_t)$
2. $\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T + R_t$
3. $K_t = \bar{\Sigma}_t H_t^T \left( H_t \bar{\Sigma}_t H_t^T + Q_t \right)^{-1}$
4. $\mu_t = \bar{\mu}_t + K_t \underbrace{(z_t - h(\bar{\mu}_t))}_{\Delta z}$ where $\Delta z$ is the innovation vector.
5. $\Sigma_t = (I - K_t H_t)\,\bar{\Sigma}_t$

   return $\mu_t, \Sigma_t$ (actually $\mathcal{N}(\mu_t, \Sigma_t)$)

Properties:

- EKF is very efficient $O(k^{2.4} + n^2)$
- Not optimal, but in practice works well (depends on the non-linearities, some are more problematic)

Compact initial distribution reduces the error because we are "near" the linearization point ($O(||\Delta x||)$)

## 2 Localization

Markov localization directly uses Bayes filter (see Figure 2):

$$\bar{bel}(x_t) = \int p(x_t | u_t, x_{t-1}, m)\; bel(x_{t-1})\; dx_{t-1}$$
$$bel(x_t) = \eta\; p(z_t | x_t, m)\; \bar{bel}(x_t)$$

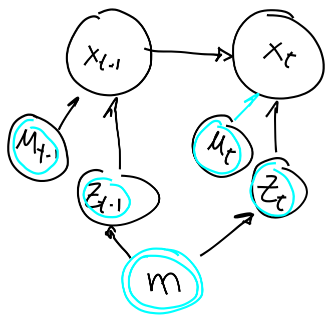


Figure 2: Bayes network corresponding to the localization problem. Observable variables are colored in cyan and estimated variables $x$'s in black.

**Example**: 1D "3 doors" problem. On Figure 3 on the 1-st plot any of the three doors could have been detected. On the 4-th plot only propagation occurs. On the 5-th plot again a door is detected, but given the previous *bel* the robot is better localizing as it is shown on the 6-th plot.

### Localization problems (taxonomy)

- Local (position tracking) [$x_0$ given] vs Global [$x_0$ unkown, kidnapped problem]
- Static vs Dynamic [moving furniture, doors, snow...]
- Passive vs Active [exploration, belief planning]
- Single-robot vs Multi-robot

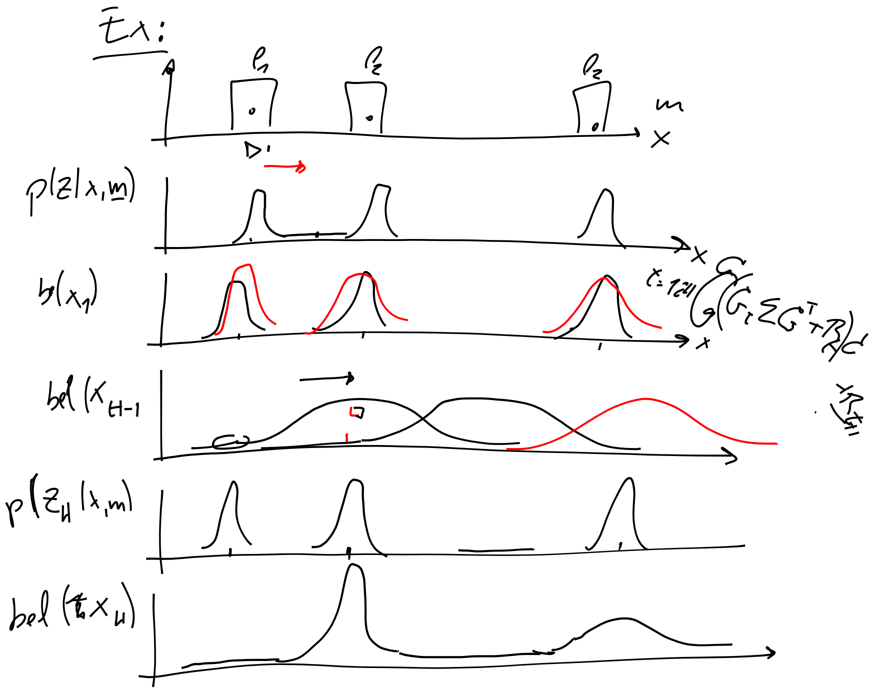


Figure 3: "3 doors" problem. In the top figure, any of the three doors could have been detected, but there is still ambiguiity in the solution, as denoted in bel$(x_1)$.

## 3 EKF localization

Gaussians are unimodal distribution and we have used 3 modes on the 3 doors problem. We need to solve the data association problem landmark-observation.

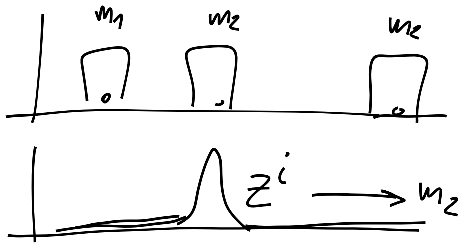


Figure 4: Unimodality of Gaussian in "3 doors" problem

We will assume known correspondences: $c^i = j$ (from landmark $m_j$). For the case of 3 observations of each landmark:

$$p(z|x, m, c) = \prod_{i=1}^{3} p(z^i|x, m, c^i)$$

### Algorithm: EKF localization with known correspondences

Inputs: $\mu_{t-1}, \Sigma_{t-1}, u_t, z_t, c_t, m$ (ProbRob 204)

1. $G_t = \frac{\partial g(x_{t-1},u_t)}{\partial x_{t-1}}$, $V_t = \frac{\partial g(x_{t-1},u_t)}{\partial u_t}$, $M_t^{arc} = \begin{bmatrix} \alpha_1 v_t^2 + \alpha_2 \omega_t^2 & 0 \\ 0 & \alpha_3 v_t^2 + \alpha_4 \omega_t^2 \end{bmatrix}$ ($M_t^{arc}$ is for arc circular model)

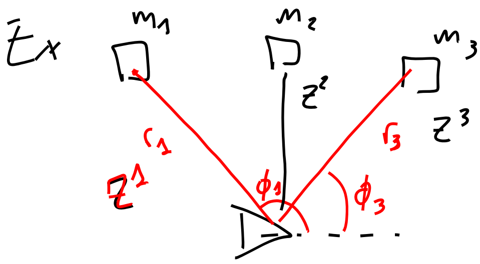


Figure 5: Data association. Now each door has a singature and we can distinguish them.

2. $\bar{\mu}_t = g(\mu_{t-1}, u_t)$
3. $\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T + V_t M_t V_t^T = G_t \Sigma_{t-1} G_t^T + R_t$
4. $Q_t = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\phi^2 \end{bmatrix}$. Range and bearing observation noise. Known correspondences $\implies$ $\sigma_s^2 = 0$ (eliminate)
5. for $\{i : z_t^i = [r_t^i, \phi_t^i]^T\}$:
6. $\quad \hat{z}_t^i = \begin{bmatrix} \sqrt{(m_{j,x} - \bar{\mu}_{t,x})^2 + (m_{j,y} - \bar{\mu}_{t,y})^2} \\ atan2(m_{j,y} - \bar{\mu}_{t,y}, m_{j,x} - \bar{\mu}_{t,x}) - \bar{\mu}_{t,\theta} \end{bmatrix}$
7. $\quad H_t^i = \frac{\partial h(x_t)}{\partial x_t}\Big|_{\bar{\mu}_t} = \begin{bmatrix} \frac{-(m_{j,x}-\bar{\mu}_{t,x})}{\sqrt{(m_{j,x}-\bar{\mu}_{t,x})^2+(m_{j,y}-\bar{\mu}_{t,y})^2}} & \frac{-(m_{j,y}-\bar{\mu}_{t,y})}{\sqrt{(m_{j,x}-\bar{\mu}_{t,x})^2+(m_{j,y}-\bar{\mu}_{t,y})^2}} & 0 \\ \frac{m_{j,y}-\bar{\mu}_{t,y}}{(m_{j,x}-\bar{\mu}_{t,x})^2+(m_{j,y}-\bar{\mu}_{t,y})^2} & \frac{-(m_{j,x}-\bar{\mu}_{t,x})}{(m_{j,x}-\bar{\mu}_{t,x})^2+(m_{j,y}-\bar{\mu}_{t,y})^2} & -1 \end{bmatrix}$
8. $\quad S_t^i = H_t^i \bar{\Sigma}_t (H_t^i)^T + Q_t$
9. $\quad K_t^i = \bar{\Sigma}_t (H_t^i)^T (S_t^i)^{-1}$
10. $\quad \bar{\mu}_t := \bar{\mu}_t + K_t^i \left(z_t^i - \hat{z}_t^i\right)$ (innovation vector for $z_t^i$)
11. $\quad \bar{\Sigma}_t := \left(I - K_t^i H_t^i\right) \bar{\Sigma}_t$
12. endfor
13. $\mu_t = \bar{\mu}_t \quad \left(\mu_t = \bar{\mu}_t + \sum_i K_t^i \left(z_t^i - \hat{z}_t^i\right)\right)$
14. $\Sigma_t = \bar{\Sigma}_t$
15. return $\mu_t, \Sigma_t$

## Iterated EKF

Why are we updating the prediction belief $\bar{bel}(x_t)$ $I$ times?
If we assume that $\{z_t^i\}_{i=1}^I$ are independent:

$$\begin{aligned} bel(x_t) =& \eta \cdot p(z|x,m,c) \cdot \bar{bel}(x_t) \\ =& \prod^I p(z^i|x,m,c) \cdot \bar{bel}(x_t) \\ =& p(z^1|x,m,c) \cdot \\ & p(z^2|x,m,c) \cdot \\ & \dots \\ & \underbrace{p(z^I|x,m,c) \cdot \bar{bel}(x_t)}_{\text{conditioning a joint Gaussian}} \end{aligned}$$

where we are iteratively conditioning a distribution, starting from $p(z^I|x, m, c) \cdot \bar{bel}(x_t)$ up to the first observation. Note that the order matters, since the linearization point at each term in the summation gets updated. Intuitively, we want them to be as close to the real solution in order to minimize the implicit linearization error.

# 4 Summary

- Extended Kalman Filter and linearization errors.
- Localization → EKF localization as a uni-modal solution.
- Map of landmarks:

$$m = \left\{ \begin{bmatrix} m_{1,x} \\ m_{1,y} \end{bmatrix} \begin{bmatrix} m_{2,x} \\ m_{2,y} \end{bmatrix} \dots \begin{bmatrix} m_{i,x} \\ m_{j,y} \end{bmatrix} \dots \right\} \text{ map of known landmark} \tag{1}$$

  The localization problem becomes a state estimation problem ⇒ EKF,UKF (in additional notes)

$$bel(x_t) = p(x_t|U, Z, m) \tag{2}$$

  Assume (for now) $c_t^i = j$ ($z^i \to m_j$) known correspondences