

L04: Bayes filter and Kalman filter

Gonzalo Ferrer

8 February 2022

1 Bayes filter general form

z : Observations \rightarrow Sensors obtain information

u : Actions \rightarrow Change the state of the world

x : State \rightarrow Robot representation and its environment

Sensor model: $p(z_t|x_t)$ — measurement probability

Action model: $p(x_t|x_{t-1}, u_t)$ — state transition probability

Belief: $bel(x_t)$ — posterior of the state

$$\text{Bel}(x_t) = p(x_t|u_1, z_1, \dots, u_t, z_t) = p(x_t|u_{1:t}, z_{1:t}) \quad (1)$$

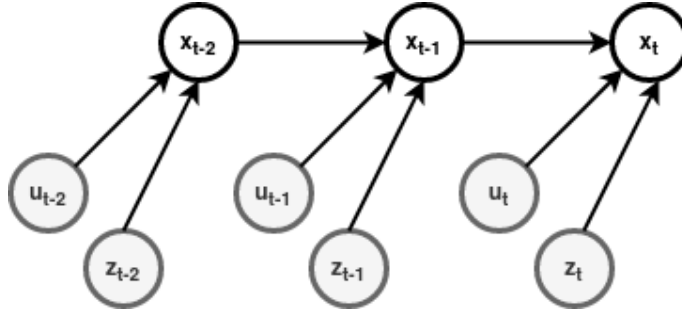


Figure 1: Graphical model

If x is complete (Markovian assumption):

$$\begin{aligned} p(x_t|x_{0:t-1}, z_{1:t-1}, u_{1:t}) &= p(x_t|x_{t-1}, u_t) \\ p(z_t|x_{0:t}, z_{1:t-1}, u_{1:t}) &= p(z_t|x_t) \end{aligned}$$

$$\begin{aligned} \text{Bel}(x_t) &= p(x_t|u_{1:t}, z_{1:t}) \\ &= \eta p(z_t|x_t, u_{1:t}, z_{1:t-1}) p(x_t|u_{1:t}, z_{t-1}) && \text{(Bayes)} \\ &= \eta p(z_t|x_t) p(x_t|u_{1:t}, z_{1:t-1}) && \text{(Markov)} \\ &= \eta p(z_t|x_t) \int_{x_{t-1}} p(x_t|u_{1:t}, z_{1:t-1}, x_{t-1}) p(x_{t-1}|u_{1:t}, z_{1:t-1}) dx_{t-1} && \text{(Total prob)} \\ &= \eta p(z_t|x_t) \int_{x_{t-1}} p(x_t|u_t, x_{t-1}) p(x_{t-1}|u_{1:t-1}, z_{1:t-1}) dx_{t-1} \\ &= \eta p(z_t|x_t) \int_{x_{t-1}} p(x_t|u_t, x_{t-1}) \text{Bel}(x_{t-1}) dx_{t-1} && \text{(Recursive form)} \end{aligned}$$

2 Bayes filter algorithm

$$\boxed{\begin{aligned} \overline{bel}(x_t) &= \int p(x_t|u_t, x_{t-1}) bel(x_{t-1}) dx_{t-1} \\ \updownarrow \\ bel(x_t) &= \eta p(z_t|x_t) \overline{bel}(x_t) \end{aligned}}$$

3 Kalman filter: Linear Dynamic system

We will derive it directly from the Bayes filter assuming a prior to be Gaussian:

$$x_{t-1} \sim \mathcal{N}(\mu_{t-1}, \Sigma_{t-1}) \quad (2)$$

3.1 (State) transition function

$$x_t = A_t x_{t-1} + B_t u_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, R) \quad (3)$$

Linear function plus added Gaussian noise $\Rightarrow x_t$ is Gaussian

$$x_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{n,t} \end{bmatrix}, \quad u_t = \begin{bmatrix} u_{1,t} \\ \vdots \\ u_{m,t} \end{bmatrix}, \quad A_t = \mathbb{R}^{n \times n}, \quad B_t = \mathbb{R}^{n \times m} \quad (4)$$

3.2 Observation function

$$z_t = C_t x_t + \delta_t, \quad \delta_t \sim \mathcal{N}(0, Q) \quad (5)$$

Linear function, δ_t is Gaussian, x_t is Gaussian $\Rightarrow z_t$ is Gaussian

$$z_t = \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ \vdots \\ z_{k,t} \end{bmatrix}, \quad C_t = \mathbb{R}^{k \times n}, \quad Q \in \mathbb{R}^{k \times k} \quad (6)$$

3.3 Kalman filter: Linear case

- Transition function is linear
- Observation function is linear
- Priors states and noise Gaussians

Then, KF (Kalman Filter) is BLUE (Best Linear Unbiased Estimator)

$$\begin{aligned} \text{I: } \bar{\mu}_t &= A_t \mu_{t-1} + B_t u_t \\ \text{II: } \bar{\Sigma}_t &= A_t \Sigma_{t-1} A_t^T + R \\ \text{III: } K_t &= \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q)^{-1} \\ \text{IV: } \mu_t &= \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t) \\ \text{V: } \Sigma_t &= (I - K_t C_t) \bar{\Sigma}_t \end{aligned}$$

Output $bel(x_t) = \mathcal{N}(\mu_t, \Sigma_t)$ becomes the input prior on the next iteration.

4 Derivation of the prediction step $\overline{bel}(x_t)$

$$\bar{\mu}_t = \mathbb{E}\{x_t | u_{1:t}, z_{1:t}\} = \mathbb{E}\{A_t x_{t-1} + B_t u_t + \varepsilon_t | \theta\} = A_t \mu_{t-1} + B_t u_t$$

In this derivation, we will use the auxiliary variable θ to express the past history of observations and odometry values. We will also include the observation z_t , although it does not affect the current transition, but it will be useful to derive the posterior distribution without any artificial assumption.

$$\begin{aligned} \bar{\Sigma}_t &= \mathbb{E}\{(x_t - \bar{\mu}_t)(x_t - \bar{\mu}_t)^T | \theta\} = \\ &= \mathbb{E}\{(A_t x_{t-1} + B_t u_t + \varepsilon_t - A_t \mu_{t-1} - B_t u_t)(A_t x_{t-1} + \varepsilon_t - A_t \mu_{t-1})^T | \theta\} \\ &= \mathbb{E}\{(A_t(x_{t-1} - \mu_{t-1}) + \varepsilon_t)(A_t(x_{t-1} - \mu_{t-1}) + \varepsilon_t)^T | \theta\} \\ &= \mathbb{E}\{A_t(x_{t-1} - \mu_{t-1})(x_{t-1} - \mu_{t-1})^T A_t^T + \varepsilon_t \varepsilon_t^T + \varepsilon_t(x_{t-1} - \mu_{t-1})^T A_t^T + \\ &\quad + A_t(x_{t-1} - \mu_{t-1})\varepsilon_t^T | \theta\} \\ &\quad \varepsilon, x - \text{uncorrelated} \quad \text{and} \quad \mathbb{E}\{\varepsilon_t\} = 0 \\ &= A_t \cdot \Sigma_{t-1} \cdot A_t^T + R \end{aligned}$$

Now we build the joint probability

$$p(x_t, x_{t-1} | \theta) = \mathcal{N} \left(\begin{bmatrix} \bar{\mu}_t \\ \mu_{t-1} \end{bmatrix}, \begin{bmatrix} A_t \Sigma_{t-1} A_t^T + R & - \\ - & \Sigma_{t-1} \end{bmatrix} \right). \quad (7)$$

Then, we marginalize this expression in order to obtain a PDF dependent only on the variable x_t :

$$\overline{bel}(x_t) = p(x_t | \theta) = \int p(x_t, x_{t-1} | \theta) dx_{t-1}.$$

Note that the result of the I and II steps coincides with the expression of the PDF we have just marginalized, for $\bar{\mu}_t$ and the block matrix in the top-left corner of the covariance above.

5 Derivation of the correction step $bel(x_t)$

Build the joint Gaussian PDF (similarly as before)

$$p(z_t, x_t | u_{1:t}, z_{1:t-1}) = p(z_t | x_t, \theta) \overline{bel}(x_t) \quad (8)$$

then condition this PDF on z_t to obtain the posterior of x_t or belief $bel(x_t)$

$$\begin{aligned} \mu_z &= \mathbb{E}\{z_t | \theta\} = \mathbb{E}\{C_t x_t + \delta_t | \theta\} = C_t \bar{\mu}_t \\ \Sigma_z &= \mathbb{E}\{(z_t - \mu_z)(z_t - \mu_z)^T | \theta\} \\ &= \mathbb{E}\{(C_t x_t + \delta_t - C_t \bar{\mu}_t)(C_t x_t + \delta_t - C_t \bar{\mu}_t)^T | \theta\} \\ &= \mathbb{E}\{C_t(x_t - \bar{\mu}_t)(x_t - \bar{\mu}_t)^T C_t^T + \delta_t(\dots)x_t^T + \delta_t \delta_t^T | \theta\} \\ &\quad (\delta_t, x_t \text{ are uncorrelated and } \mathbb{E}\{\delta_t\} = 0.) \\ &= C_t \bar{\Sigma}_t C_t^T + Q \\ \Sigma_{x,z} &= cov(x_t, z_t) = \mathbb{E}\{(x_t - \bar{\mu}_t)(z_t - \bar{\mu}_z)^T | \theta\} \\ &= \mathbb{E}\{(x_t - \bar{\mu}_t)(C_t x_t + \delta_t - C_t \bar{\mu}_t)^T | \theta\} = \\ &= \mathbb{E}\{(x_t - \bar{\mu}_t)(x_t - \bar{\mu}_t)^T C_t^T + x_t \delta_t - \bar{\mu}_t \delta_t^T | \theta\} = \bar{\Sigma}_t C_t^T \end{aligned}$$

$$p(x_t, z_t | \mu_{1:t}, z_{1:t-1}) = N \left(\begin{bmatrix} C_t \bar{\mu}_t \\ \bar{\mu}_t \end{bmatrix}, \begin{bmatrix} C_t \bar{\Sigma}_t C_t^T + Q & C_t \bar{\Sigma}_t \\ \bar{\Sigma}_t C_t^T & \bar{\Sigma}_t \end{bmatrix} \right)$$

$$\begin{aligned} p(x_t | z_t, \mu_{1:t}, z_{1:t-1}) &= bel(x_t) = \mathcal{N}(\bar{\mu}_t + \Sigma_{z,x} \Sigma_z^{-1} (z_t - \mu_z), \bar{\Sigma}_t - \Sigma_{x,z} \Sigma_z^{-1} \Sigma_{z,x}) \\ \mu_t &= \bar{\mu}_t + \underbrace{\bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q)^{-1} (z_t - C_t \bar{\mu}_t)}_{K_t} \quad (\text{III}) \\ &= \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t) \quad (\text{IV}) \\ \Sigma_t &= \bar{\Sigma}_t - \underbrace{\bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q)^{-1} C_t \bar{\Sigma}_t}_{K_t} = (I - K_t C_t) \bar{\Sigma}_t \quad (\text{V}) \end{aligned}$$

KF

- Highly efficient $O(k^3 + n^2)$
- Optimal for linear Gaussian systems
- Most real world system are non-linear.

6 Summary

In this lecture, we have described the following:

Bayes filter

$$\begin{aligned} \overline{bel}(x_t) &= \int p(x_t | x_{t-1}, u_t) bel(x_{t-1}) dx_{t-1} \\ bel(x_t) &= \eta p(z_t | x_t) \overline{bel}(x_t) \end{aligned}$$

Kalman filter: Bayes filter for linear system and Gaussians

prediction (Steps I, II)	Marginalization G 's
correction (Steps III, IV, V)	Conditioning G 's

$$\begin{cases} x_t &= g(x_{t-1}, u_t, \varepsilon_t) = A_t x_{t-1} + B_t u_t + \varepsilon_t \\ z_t &= h(x_t, \delta_t) = C_t x_t + \delta_t \end{cases}$$