

L04: Bayes filter and Kalman filter

Summary from L02, L03:

- Expectation: manipulations, mean, covariance
- $N(x; \mu, \Sigma)$
- Covariance projection: linear, affine and Non-linear.
- Gaussian visualization (iscontours) \Downarrow still Gaussians
- Sampling from Gaussians
- Marginalization and conditioning on Joint Gaussian.

$$p(x_a, x_b) = N\left(\begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix}, \begin{bmatrix} \Sigma_a & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_b \end{bmatrix}\right)$$

* 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.

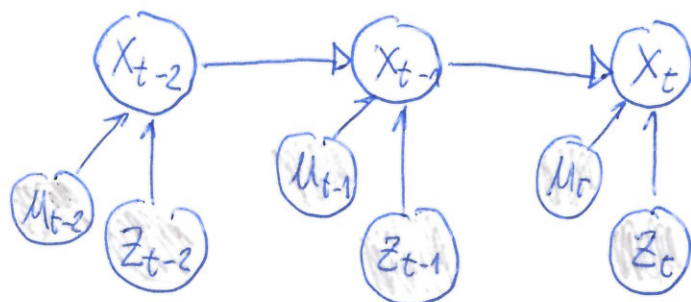
Sensor model $p(z_t | x_t)$ measurement probability.

Action model $p(x_t | x_{t-1}, u_t)$ State transition probability.

Belief: posterior of the state.

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

Graphical model:



if x is complete (Markovian assumption):

$$p(x_t | x_{0:t-1}, z_{1:t-1}, u_{1:t}) = p(x_t | x_{t-1}, u_t)$$

$$p(z_t | x_{0:t-1}, z_{1:t-1}, u_{1:t}) = p(z_t | x_t)$$

$$\bullet \text{ Bel}(x_t) = p(x_t | u_{1:t}, z_{1:t})$$

$$\text{(Bayes)} \quad = \eta \, p(z_t | x_t, u_{1:t}, z_{1:t-1}) p(x_t | u_{1:t}, z_{t-1})$$

$$\text{(Markov)} \quad = \eta \, p(z_t | x_t) p(x_t | u_{1:t}, z_{1:t-1})$$

$$\text{(total prob)} \quad = \eta \, p(z_t | x_t) \int_{x_{t-1}} p(x_t | u_{1:t}, z_{1:t-1}, x_{t-1}) p(x_{t-1} | u_{1:t}, z_{1:t-1}) dx_{t-1}$$

$$= \eta \, p(z_t | x_t) \int_{x_{t-1}} p(x_t | u_t, x_{t-1}) p(x_{t-1} | u_{1:t-1}, z_{1:t-1}) dx_{t-1}$$

$$= \eta \, p(z_t | x_t) \int_{x_{t-1}} p(x_t | u_t, x_{t-1}) \text{Bel}(x_{t-1}) dx_{t-1}$$

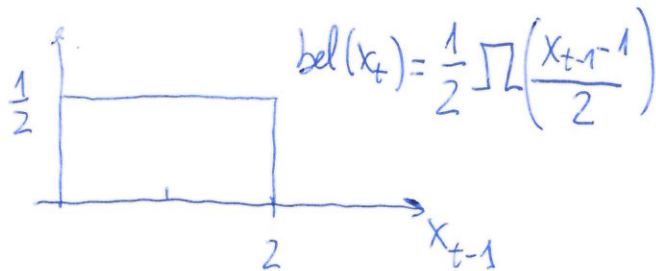
(Recursive form)

* Bayes filter algorithm

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

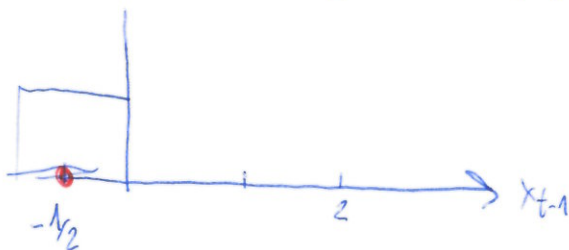
$$bel(x_t) = \eta p(z_t | x_t) \bar{bel}(x_t)$$

Ex: 1d robot, no observations



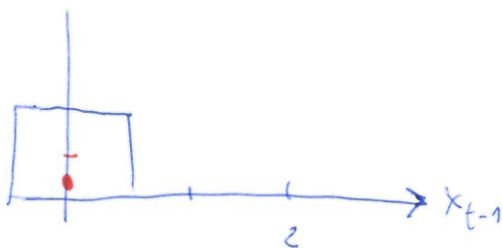
$$x_t = x_{t-1} + u_t$$

$$p(x_t | u_t, x_{t-1}) = \mathbb{I}(x_t - x_{t-1} - u_t)$$

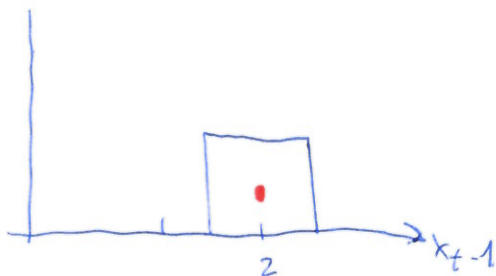


$$u=0$$

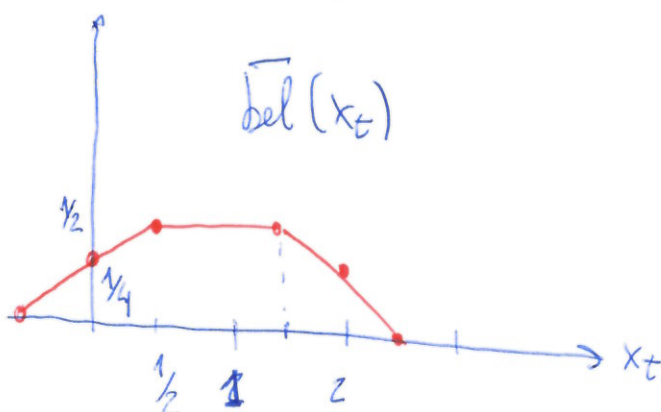
$$x_t = -\frac{1}{2}$$



$$x_t = 0$$



$$x_t = 2$$



$$\bar{bel}(x_t) = \int p(x_t | x_{t-1}, u) \bar{bel}(x_{t-1}) dx_{t-1}$$

Q: next iteration?

* Kalman filter; Dynamic system.

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

$$X_{t-1} \sim N(\mu_{t-1}, \Sigma_{t-1})$$

o (State) transition function

$$X_t = A_t X_{t-1} + B_t u_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, R)$$

Linear function plus added Gaussian noise $\Rightarrow X_t$ Gaussian

$$X_t = \begin{bmatrix} X_{1,t} \\ X_{2,t} \\ \vdots \\ X_{n,t} \end{bmatrix}, \quad u_t = \begin{bmatrix} u_{1,t} \\ \vdots \\ u_{m,t} \end{bmatrix}, \quad A_t = \begin{bmatrix} \mathbb{R}^{n \times n} \end{bmatrix} \begin{matrix} \uparrow \\ n \\ \leftarrow n \rightarrow \end{matrix}, \quad B = \begin{bmatrix} \mathbb{R}^{n \times m} \end{bmatrix} \begin{matrix} \uparrow \\ n \\ \leftarrow m \rightarrow \end{matrix}$$

(rows x columns)

o Observation function

$$Z_t = C_t X_t + S_t, \quad S_t \sim N(0, Q)$$

Linear function, S_t Gaussian, X_t Gaussian $\Rightarrow Z_t$ Gaussian.

$$Z_t = \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ \vdots \\ Z_{k,t} \end{bmatrix}, \quad C_t = \begin{bmatrix} \mathbb{R}^{k \times n} \end{bmatrix} \begin{matrix} \uparrow \\ k \\ \leftarrow n \rightarrow \end{matrix}, \quad Q \in \mathbb{R}^{k \times k}$$

Ex: Temperature system:

x : temperature [$^{\circ}\text{C}$] $u = \begin{cases} \text{'heating on'} \\ \text{'heating off'} \end{cases}$

$$x_t = A \cdot x_{t-1} + B u_t + \varepsilon_t \quad (A=0.95)$$

$$z_t = C x_t + \delta_t \quad z: \text{'Volts from sensor'}$$

[V/ $^{\circ}\text{C}$]

* Kalman filter: Linear case

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

Then, KF is BLUE (Best Linear Unbiased Estimator)

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

II: $\bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R$

III: $K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q)^{-1}$

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

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

Prediction

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

Correction/
Update.

$$\text{bel}(x_t) = \gamma p(z_t | x_t) \bar{\text{bel}}(x_t)$$

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

Derive.

* Show for prediction.

$$\bar{\mu}_t = E\{ \bar{x}_t \} = E\{ A_t x_{t-1} + B_t u_t + \varepsilon_t \} = A_t \mu_{t-1} + B_t u_t$$

$$\bar{\Sigma}_t = E\{ (\bar{x}_t - \bar{\mu}_t)(\bar{x}_t - \bar{\mu}_t)^T \}$$

$$= E\{ (A_t x_{t-1} + B_t u_t + \varepsilon_t - A_t \mu_{t-1} - B_t u_t)(A_t x_{t-1} + \varepsilon_t - A_t \mu_{t-1})^T \}$$

$$= E\{ (A_t (x_{t-1} - \mu_{t-1}) + \varepsilon_t)(A_t (x_{t-1} - \mu_{t-1}) + \varepsilon_t)^T \}$$

$$= E\{ A_t (x_{t-1} - \mu_{t-1})(x_{t-1} - \mu_{t-1})^T A_t^T + \varepsilon_t \varepsilon_t^T +$$

$$\cancel{\varepsilon_t (x_{t-1} - \mu_{t-1})^T A_t^T} + \cancel{A_t (x_{t-1} - \mu_{t-1}) \varepsilon_t^T} \}$$

ε, x uncorrelated $E\{ \varepsilon_t \} = 0$

$$= A_t E\{ \underbrace{(x_{t-1} - \mu_{t-1})(x_{t-1} - \mu_{t-1})^T}_{\bar{\Sigma}_{t-1}} \} A_t^T + R$$

Now we build the joint probability $p(x_t, x_{t-1})$ as a joint Gaussian.

$$p(\bar{x}_t, x_{t-1}) = N \left(\begin{bmatrix} \bar{\mu}_t \\ \mu_{t-1} \end{bmatrix}, \begin{bmatrix} \text{I} & \text{II} \\ \text{II}^T & \bar{\Sigma}_{t-1} \end{bmatrix} \right)$$

(I) (II)

and marginalize $p(\bar{x}_t) = \int p(\bar{x}_t, x_{t-1}) dx_{t-1}$.

* Derive for correction step

Build $p(z_t, \bar{x}_t) = p(z_t | \bar{x}_t) \text{bel}(x_t)$,

then we condition on z_t to obtain $\text{bel}(x_t)$

$$\mu_z = E\{z_t\} = E\{C_t \bar{x}_t + \delta_t\} = C_t \bar{\mu}_t$$

$$\Sigma_z = E\{(z_t - \mu_z)(z_t - \mu_z)^T\} =$$

$$= E\{(C_t \bar{x}_t + \delta_t - C_t \bar{\mu}_t)(C_t \bar{x}_t + \delta_t - C_t \bar{\mu}_t)^T\} =$$

$$= E\{C_t (\bar{x}_t - \bar{\mu}_t)(\bar{x}_t - \bar{\mu}_t)^T C_t^T + \cancel{\delta_t \bar{x}_t} + \delta_t \delta_t^T\} =$$

$$= C_t \bar{\Sigma}_t C_t^T + Q \quad \delta, \bar{x} \text{ uncorrelated} + E\{\delta\} = 0$$

$$\Sigma_{\bar{x}, z} = \text{cov}(\bar{x}_t, z_t) = E\{(\bar{x}_t - \bar{\mu}_t)(z_t - \mu_z)^T\} =$$

$$= E\{(\bar{x}_t - \bar{\mu}_t)(C_t \bar{x}_t + \delta_t - C_t \bar{\mu}_t)^T\} =$$

$$= E\{(\bar{x}_t - \bar{\mu}_t)(\bar{x}_t - \bar{\mu}_t)^T C_t^T + \cancel{\bar{x}_t \delta_t} + \cancel{\bar{\mu}_t \delta_t}\} =$$

$$= \bar{\Sigma}_t C_t^T$$

$$\left(\Sigma_{z, \bar{x}} = \Sigma_{\bar{x}, z}^T \right) \quad (4.8)$$

$$p(\bar{x}_t, z_t) = N \left(\begin{bmatrix} C_t \bar{\mu}_t \\ \bar{\mu}_t \end{bmatrix}, \begin{bmatrix} C_t \bar{\Sigma}_t C_t^T + Q & C_t^T \bar{\Sigma}_t \\ \bar{\Sigma}_t C_t^T & \bar{\Sigma}_t \end{bmatrix} \right)_{(\bar{x}_t, z_t)} \quad \text{from (4.3)}$$

$$p(x_t | z_t) = N \left(\bar{\mu}_t + \Sigma_{x, z} \Sigma_z^{-1} (z_t - \mu_z), \bar{\Sigma}_t - \Sigma_{x, z} \Sigma_z^{-1} \Sigma_{z, x} \right) \\ = N(\mu_t, \Sigma_t) = \text{bel}(x_t)$$

$$\mu_t = \bar{\mu}_t + \underbrace{\bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q)^{-1}}_{K_t} (z_t - C_t \bar{\mu}_t) \quad \textcircled{\text{III}}$$

$$= \bar{\mu}_t + K_t \underbrace{(z_t - C_t \bar{\mu}_t)}_{\text{IV}} \quad \textcircled{\text{IV}}$$

$$\Sigma_t = \bar{\Sigma}_t - \underbrace{\bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q)^{-1} C_t \bar{\Sigma}_t}_{K_t}$$

$$= \underbrace{(\mathbf{I} - K_t C_t) \bar{\Sigma}_t}_{\text{V}} \quad \textcircled{\text{V}}$$

- KF:
- Highly efficient $O(K^{2.376} + n^2)$
 - Optimal for linear Gaussian systems
 - Most robotics/real world systems are non-linear...

Ex: Temperature system.

$$x_0 \sim N(100, 10)$$

$$\bar{x}_t = 0.9 x_{t-1} + 0.1 u_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, 1)$$

$$z_t = 0.3 \bar{x}_t + d_t, \quad d_t \sim N(0, 4)$$

$A, B, C, ?$

if $u_1 = 0$, $\bar{\mu}_1?$ $\bar{\Sigma}_1?$

$$\bar{\mu}_1 = A_t \mu_{t-1} + B_t u_t = 0.9 \cdot 100 + 0.1 \cdot 0 = 90$$

$$\bar{\Sigma}_1 = A_t \Sigma_{t-1} A_t + R = 0.9^2 \cdot 10 + 1 = 9.1$$

if $z=30$, $K_1?$ $x_1?$ $\Sigma_1?$

$$K_1 = \bar{\Sigma}_t C_t (C_t \bar{\Sigma}_t C_t + Q)^{-1}$$

$$= \underbrace{9.1 \cdot 0.3}_{2.73} \left(\underbrace{0.3^2 \cdot 9.1}_{0.819} + \underbrace{4}_{\text{dominates}} \right)^{-1} = 0.5665$$

($Q=0 \Rightarrow K_t=3.33$)
 $(4.819)^{-1} \approx 0.2$

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

$$= 90 + 0.5665 \left(30 - \underbrace{0.3 \cdot 90}_{27} \right) = 91.665$$

$30/0.3 = 100$

$$\left(Q=0 \Rightarrow \mu_t = 90 + 3.33 \cdot 3 = 100. \right) \text{ perfect sensing.}$$

$$\Sigma_1 = (I - K_1 C_1) \bar{\Sigma}_1 = \underbrace{\left(1 - 0.5665 \cdot 0.3 \right)}_{0.8} 9.1 = 7.55$$

0.8: Contracting term.

4.10

