L04: Bayes filter and Kalman filter

Summary from L02, L03:

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

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

* 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.

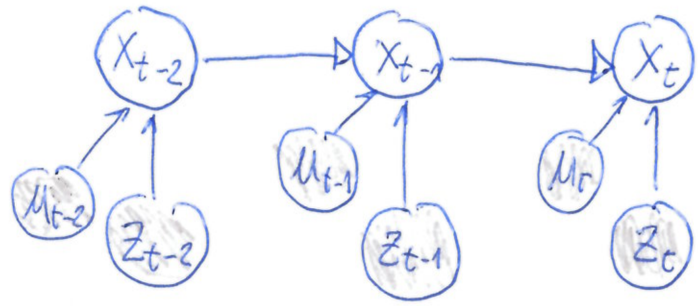
Sensor model $p(z_t | x_t)$ measurement probability.

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

Belief: posterior of the state.

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

Graphical model:



if x is complete (Markovian assumption):

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

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

- $Bel(x_t) = p(x_t \mid u_{1:t}, z_{1:t})$

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

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

(Total prob) $$= \eta\, p(z_t \mid x_t) \int_{x_{t-1}} p(x_t \mid u_{1:t}, z_{1:t-1}, x_{t-1})\, p(x_{t-1} \mid u_{1:t}, z_{1:t-1}) \cdot dx_{t-1}$$

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

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

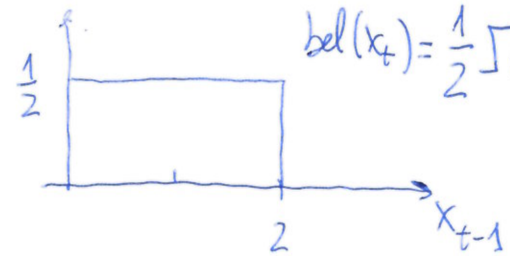
(Recursive form)

* Bayes filter algorithm

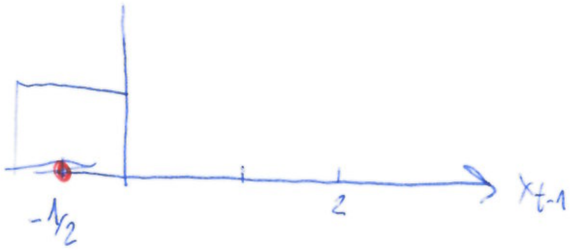
$$\overline{bel}(x_t) = \int p(x_t \mid u_t, x_{t-1})\, bel(x_{t-1})\, dx_{t-1}$$
$$bel(x_t) = \eta\, p(z_t \mid x_t)\, \overline{bel}(x_t)$$

Ex: 1d. robot, no observations

$$bel(x_t) = \frac{1}{2} \Pi\left(\frac{x_{t-1}-1}{2}\right)$$

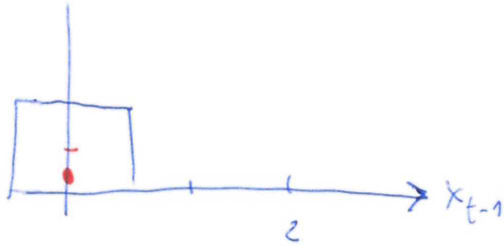


$$x_t = x_{t-1} + u_t$$
$$p(x_t \mid u_t, x_{t-1}) = \Pi(x_t - x_{t-1} - u_t)$$

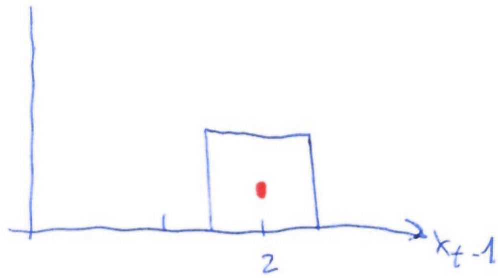


$u = 0$

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

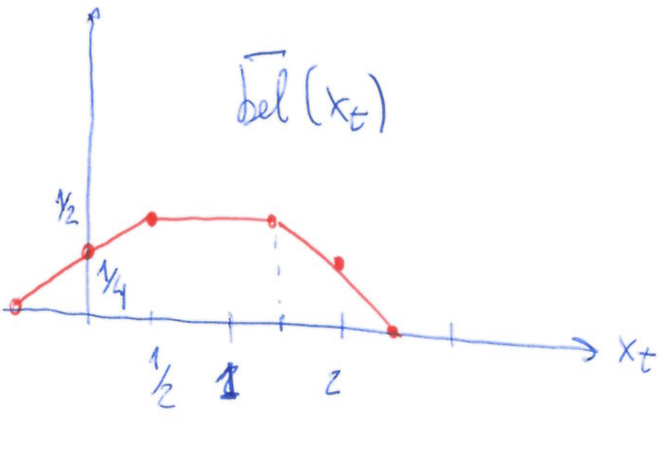


$x_t = 0$



$x_t = 2$

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



Q: next iteration?

* Kalman filter; Dynamic system.

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

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

- (State) transition function

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

Linear function pluss added Gaussian noise $\Rightarrow x_t$ Gaussian

$$x_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{n,t} \end{bmatrix}, \quad u_t = \begin{bmatrix} u_{1,t} \\ \vdots \\ u_{m,t} \end{bmatrix} \quad A_t = \underbrace{\begin{bmatrix} & & \\ & \mathbb{R}^{n\times n} & \\ & & \end{bmatrix}}_{n} \Big\updownarrow n, \quad B = \underbrace{\begin{bmatrix} \mathbb{R}^{n\times m} \end{bmatrix}}_{m} \Big\updownarrow n$$

(rows x columns)

- Observation function

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

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

$$z_t = \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ \vdots \\ z_{k,t} \end{bmatrix}, \quad C_t = \underbrace{\begin{bmatrix} & & \\ & \mathbb{R}^{k\times n} & \\ & & \end{bmatrix}}_{n} \Big\updownarrow k, \quad Q \in \mathbb{R}^{k\times k}$$

Ex: Temperature system:

x : temperature [°C] $\quad u = \{$'heating on', 'heating off'$\}$

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

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

$[V/°C]$

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

Prediction

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

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

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

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

Correction / Update.

$bel(x_t) = \eta \, p(z | x) \, \overline{bel}(x_t)$

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

* Derive prediction $\overline{bel}(x_t)$

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

$$\bar{\Sigma}_t = E\{(x_t - \bar{\mu}_t)(x_t - \bar{\mu}_t) \mid \theta\}$$

we will omit this later, but it must considered

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

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

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

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

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

$$= A_t \cdot \Sigma_{t-1} \cdot A_t^T + R$$

Now we build the joint probability $p(x_t, x_{t-1} \mid \theta)$

$$p(x_t, x_{t-1} \mid \theta) = \mathcal{N}\left(\begin{bmatrix} \boxed{\bar{\mu}_t}^{\,\mathrm{I}} \\ \mu_{t-1} \end{bmatrix}, \begin{bmatrix} \boxed{A_t \Sigma_{t-1} A_t^T + R}^{\,\mathrm{II}} & - \\ - & \Sigma_{t-1} \end{bmatrix} \right)$$

marginalize $\quad p(x_t \mid \theta) = \int p(x_t, x_{t-1} \mid \theta)\, dx_{t-1}$

$$\| $$

$$\overline{bel}(x_t)$$

* Derive for correction step $bel(x_t)$

Build $p(z_t, x_t \mid \underbrace{u_{1:t}, z_{1:t-1}}_{\theta}) = p(z_t \mid x_t, \theta)\, \overline{bel}(x_t)$

then, condition on z_t to obtain $bel(x_t)$

$$\mu_z = E\{z_t \mid \theta\} = E\{C_t x_t + \delta_t \mid \theta\} = C_t \bar{\mu}_t$$

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

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

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

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

δ, x uncorrelated

$E\{\delta\} = 0$

$$\Sigma_{x,z} = cov(x_t, z_t) = E\{(x_t - \mu_t)(z_t - \mu_z) \mid \theta\} =$$

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

$$= E\{(x_t - \bar{\mu}_t)(x_t - \bar{\mu}_t)^T C^T + x_t \delta_t - \mu_t \delta_t \mid \theta\}$$

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

$$p(x_t, z_t \mid u_{1:t}, z_{1:t-1}) = \mathcal{N}\left(\begin{bmatrix} C_t \bar{\mu}_t \\ \bar{\mu}_t \end{bmatrix}, \begin{bmatrix} C_t \bar{\Sigma}_t C_t^T + Q & \overset{(\Sigma_{z,x})}{C_t \bar{\Sigma}_t} \\ \bar{\Sigma}_t C_t^T & \bar{\Sigma}_t \end{bmatrix} \right)$$

joint Gaussian conditioning.

$$(\Sigma_{x,z} = \Sigma_{z,x}^T)$$

$$p(x_t \mid z_t, u_{1:t}, z_{1:t-1}) = bel(x_t) =$$

$$= \mathcal{N}\left(\bar{\mu}_t + \Sigma_{x,z} \Sigma_z^{-1} (z_t - \mu_z), \bar{\Sigma}_t - \Sigma_{x,z} \Sigma_z^{-1} \Sigma_{z,x} \right)$$

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

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

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

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

KF: - Highly efficient $O(k^3 + n^2)$ k: obs dim.

- Optimal for linear Gaussian systems
- Most real world systems are non-linear.

Ex: Temperature system. $x_0 \sim N(100, 10)$

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

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

$A, B, C, ?$

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? Σ_1?

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

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

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

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

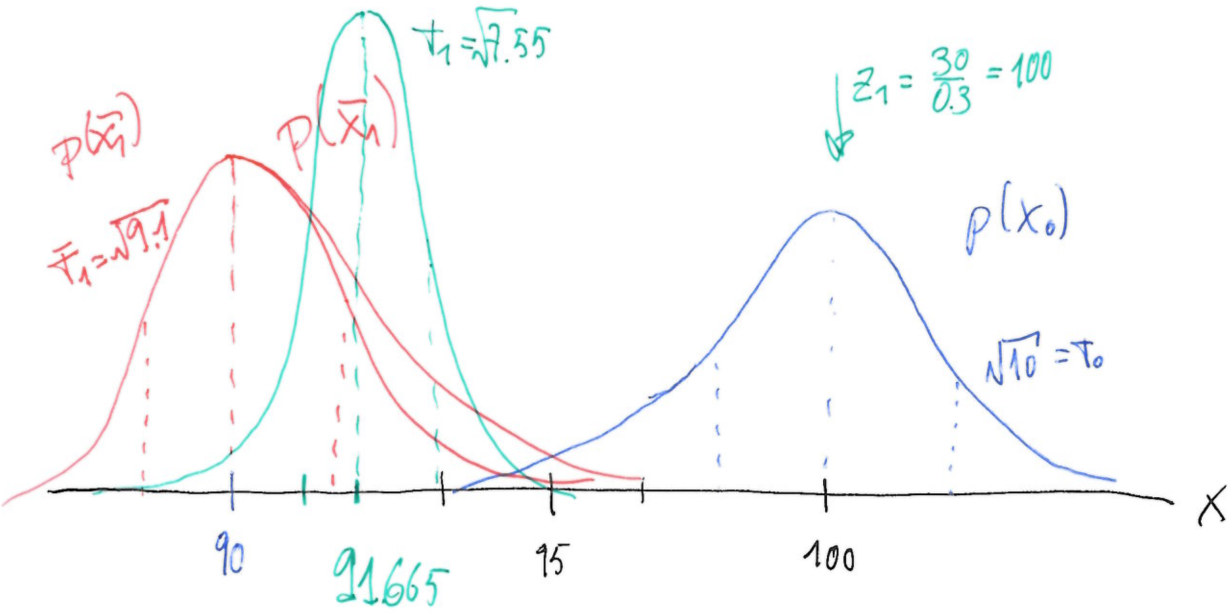
$$= 90 + 0.5665 \Big(30 - \underbrace{0.3 \cdot 90}_{27} \Big) = 91.665$$

$30/0.3 = 100$

$(Q = 0 \Rightarrow \mu_t = 90 + 3.33 \cdot 3 = 100.)$ perfect sensing.

$$\Sigma_1 = (I - K_1 C_1) \bar{\Sigma}_1 = \Big(\underbrace{1 - 0.5665 \cdot 0.3}_{0.8 \text{ contracting term.}} \Big) 9.1 = 7.55$$

4.10



$p(\bar{x}_1)$
$\sigma_1 = \sqrt{9.1}$
$p(\bar{x}_1)$
$\sigma_1 = \sqrt{7.55}$
$z_1 = \frac{30}{0.3} = 100$
$p(x_0)$
$\sqrt{10} = \sigma_0$
90
91.665
95
100
X