Kalman filter

Gellért Peresztegi-Nagy

April 16, 2017

Question 1.

The following graph shows the graphical model of the Kalman filter, visualizing the dependencies between the variables:

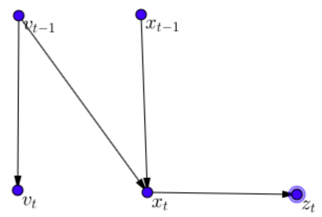

Figure 1: The graphical model of the Kalman filter

Question 2.

If we have $v_{t-1} = \mathcal{N}(v, \sigma^2_{v_{prior}})$ and $x_{t-1} = \mathcal{N}(x, \sigma^2_{x_{prior}})$, $v_t = v_{t-1} + \epsilon$ where $\epsilon = \mathcal{N}(0, \sigma^2_v)$.

As the sum of two normal distributions is a normal distribution as well:

$$v_t = \mathcal{N}(v, \sigma^2_{v_{prior}}) + \mathcal{N}(0, \sigma^2_v) = \mathcal{N}(v, \sigma^2_{v_{prior}} + \sigma^2_v)$$

Question 3.

Similarly:

$$x_t = x_{t-1} + v_{t-1} = \mathcal{N}(x, \sigma^2_{x_{prior}}) + \mathcal{N}(v, \sigma^2_{v_{prior}}) = \mathcal{N}(x + v, \sigma^2_{x_{prior}} + \sigma^2_{v_{prior}})$$

Question 4.

We can assume that $P(x_{t-1}|z_{t-1}, z_{t-2}, \ldots, z_1) = \mathcal{N}(\mu_x, \sigma^2_x)$ and that $P(v_{t-2}|z_{t-1}, z_{t-2}, \ldots, z_1) = \mathcal{N}(\mu_v, \sigma^2_v)$. We would like to get the posterior of x_t having seen z_t as well.

Using Bayes' rule, we know that:

$$P(x_t|z_t, z_{t-1}, z_{t-2}, \ldots, z_1) \propto P(z_t|x_t, z_{t-1}, z_{t-2}, \ldots, z_1)P(x_t|z_{t-1}, z_{t-2}, \ldots, z_1)$$

From the graphical model (Figure 1), we can see that $P(z_t|x_t, z_{t-1}, z_{t-2}, \dots, z_1) = P(z_t|x_t)$ as there is only one edge going to z_t which is from x_t, hence knowing x_t makes the route blocked (d-separation). Therefore, knowing x_t makes z_t conditionally independent of $z_{t-1}, \dots, z_1$. Hence:

$$P(x_t|z_t, z_{t-1}, z_{t-2}, \dots, z_1) \propto P(z_t|x_t)P(x_t|z_{t-1}, z_{t-2}, \dots, z_1)$$

z_t is a normal distribution around x_t with standard deviation σ_z. However, because of the properties of the normal distribution ($(\mu - x)^2 = (x - \mu)^2$), (knowing z_t) we can think of x_t as a normal distribution around z_t with σ_z standard deviation. (This is indeed logical as knowing z_t suggests that x_t must be around z_t). Hence: $P(x_t|z_t) = \mathcal{N}(z_t, \sigma_z^2)$.

As $x_t = x_{t-1} + v_{t-1} = x_{t-1} + v_{t-2} + \epsilon$ where $\epsilon = \mathcal{N}(0, \sigma_{v_0}^2)$,
$P(x_t|z_{t-1}, z_{t-2}, \dots, z_1) \propto \mathcal{N}(\mu_{x_{t-1}} + \mu_{v_{t-2}}, \sigma_x^2 + \sigma_v^2 + \sigma_{v_0}^2)$.

$$P(x_t|z_t, z_{t-1}, z_{t-2}, \dots, z_1) \propto \mathcal{N}(\mu_{x_{t-1}} + \mu_{v_{t-2}}, \sigma_{x_{t-1}}^2 + \sigma_{v_{t-2}}^2 + \sigma_{v_0}^2) \cdot \mathcal{N}(\mu_{x_{t-1}}, \sigma_{x_t}^2)$$

By ignoring the factors not depending on x_t and rearranging the normal distributions, we get that:

$$P(x_t|z_t, z_{t-1}, z_{t-2}, \dots, z_1) \sim \mathcal{N}(\mu_{x_t}, \sigma_{x_t}^2)$$

Where $\mu_{x_t} = \frac{\frac{z_t}{\sigma_z^2} + \frac{\mu_{x_{t-1}} + \mu_{v_{t-2}}}{\sigma_{x_{t-1}}^2 + \sigma_{v_{t-2}}^2 + \sigma_{v_0}^2}}{\sigma_z^{-2} + (\sigma_{x_{t-1}}^2 + \sigma_{v_{t-2}}^2 + \sigma_{v_0}^2)^{-1}}$ and $\sigma_{x_t}^{-2} = \sigma_z^{-2} + (\sigma_{x_{t-1}}^2 + \sigma_{v_{t-2}}^2 + \sigma_{v_0}^2)^{-1}$

Question 5.

Having the same assumptions about x_t and v_{t-1} as in Question 4 and writing $v_{t-1} = x_t - x_{t-1}$ we can obtain an expression for the posterior distribution $P(v_{t-1}|z_t, z_{t-1}, \dots, z_1)$

From Question 4 we know that $P(v_{t-1}|z_t, z_{t-1}, z_{t-2}, \dots, z_1) \sim \mathcal{N}(\mu_{x_t}, \sigma_{x_t}^2) - \mathcal{N}(\mu_{x_{t-1}}, \sigma_{x_{t-1}}^2)$. Hence, $P(v_{t-1}|z_t, z_{t-1}, z_{t-2}, \dots, z_1) \sim \mathcal{N}(\mu_{v_{t-1}}, \sigma_{v_{t-1}}^2)$ where $\mu_{v_{t-1}} = \mu_{x_t} - \mu_{x_{t-1}}$ and $\sigma_{v_{t-1}}^2 = \sigma_{x_t}^2 + \sigma_{x_{t-1}}^2$.