Control theory - part 3, state estimation, Kalman filter

Marek Cygan

January 13, 2024

1 Intro

Consider a random variable x that models our belief about the state. Imagine we receive a sensory measurement z, that follows a distribution $p(z|x)$ - here we condition on the state as in the causal part of the sensory model this conditional distribution is the easiest to model.

Our goal is to derive the conditional probability $p(x|z)$, i.e., to update our belief about the world by taking into account the measurment z.

Assumptions:

- $p(x) \sim N(\mu_x, \sigma_x^2)$ - here we assume that our prior is a Gaussian,
- $p(z|x) \sim N(x, \sigma_z^2)$ - when conditioned on x, the random variable z is a Gaussian centered around x with a variance σ_z^2.

In what follows we show that $p(x|z)$ (i.e., the updated belief, also called posterior) is still a Gaussian.

Bayes' theorem states:

$$p(x|z) = \frac{p(z|x) \cdot p(x)}{p(z)}$$

Since $p(z)$ is just normalizing, it is enough to show that $p(z|x) \cdot p(x)$ follows a Gaussian distribution, which we do by inspecting the product of the pdfs of the two distributions. Since $x \sim \mathcal{N}(\mu_x, \sigma^2)$ and $z \sim \mathcal{N}(x, \sigma_z^2)$, the probability density functions (PDFs) for these distributions are:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right)$$

$$p(z|x) = \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left(-\frac{(z-x)^2}{2\sigma_z^2}\right)$$

Plugging these into Bayes' theorem and skipping the constants that only normalize, we get:

$$\begin{aligned} p(x|z) &\sim \exp\left(-\frac{(z-x)^2}{2\sigma_z^2}\right) \cdot \exp\left(-\frac{(x-\mu_x)^2}{2\sigma^2}\right) \\ &= \exp\left(-\frac{1}{2}\left(\frac{(x-z)^2}{\sigma_z^2} + \frac{(x-\mu_x)^2}{\sigma^2}\right)\right) \end{aligned} \tag{1}$$

In order to express this function in the form $c \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, we use the following regrouping technique:

$$\begin{aligned} ax^2 + bx + c &= (\sqrt{a}x)^2 + 2 \cdot \sqrt{a}x \cdot \frac{b}{2\sqrt{a}} + \frac{b^2}{4a} + (c - \frac{b^2}{4a}) \\ &= \frac{(x + \frac{b}{2a})^2}{1/a} + (c - \frac{b^2}{4a}) \end{aligned} \tag{2}$$

Applying (??) to (??) gives $a = \sigma_z^{-2} + \sigma_x^{-2} = \frac{\sigma_x^2 + \sigma_z^2}{\sigma_x^2 \sigma_z^2}$, $b = \frac{-2z}{\sigma_z^2} + \frac{-2\mu_x}{\sigma_x^2} = \frac{-2z\sigma_x^2 - 2\mu_x\sigma_z^2}{\sigma_z^2\sigma_x^2}$ which leads to

$$
\begin{aligned}
p(x|z) &\sim \exp\left(-\frac{1}{2}\left(\frac{(x-z)^2}{\sigma_z^2}+\frac{(x-\mu_x)^2}{\sigma_x^2}\right)\right)\\
&\sim \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \text{ where}\\
\mu &=\frac{\sigma_z^2\mu+\sigma_x^2 z}{\sigma_x^2+\sigma_z^2}\\
\sigma^2 &=\frac{\sigma_x^2\sigma_z^2}{\sigma_x^2+\sigma_z^2}
\end{aligned}
\tag{3}
$$

2 Multivariate case

We use the property of multivariate normal distributions, that they are closed under linear transformations.

2.1 Preliminaries

If a multivariate normal random variable $X \sim \mathcal{N}(\mu_X, \Sigma_X)$ is multiplied by a constant matrix A, the resulting random variable AX is also normally distributed. The new distribution is given by:

$$AX \sim \mathcal{N}(A\mu_X, A\Sigma_X A^T)$$

2.2 Updating the posterior given a sensory reading

In the following theorem we assume that the sensory measurement is a noisy projection.

Theorem 1. *If $p(x) \sim N(\mu, \Sigma)$ and $p(z|x) \sim N(Hx, R)$ then*

$$p(z) \sim N(H\mu, S) \textit{ where } S = H\Sigma H^T + R$$

and

$$p(x|z) \sim N(\mu + K(y - H\mu), \Sigma - KH\Sigma), \textit{ where } K = \Sigma H^T S^{-1}$$

Note that in the theorem above $H\mu$ is the mean of $p(z|x)$, which is why $z - H\mu$ is called a residual.