INFO8006 Introduction to Artificial Intelligence

Kalman Filter

Exercise 1: Hyperloop

This year ULiège has decided to get into the Hyperloop competition (`https://www.spacex.com/hyperloop`). Briefly, what you should do to win this competition is to build the fastest and most reliable autonomous pod. One of the most important engineering problem to build the pod is to be able to compute a robust estimation of the state of the pod (its position and speed) given many noisy sensors.

This morning you received an email asking you what would be your solution to this estimation problem. The message contains information about the sensors they plan to put in the pod. They say that they will use 3 unbiased speed sensors with a 99.7% accuracy[1] of 0.1m/s and a GPS sensor (also unbiased) which provides the pod position (in one dimension) with a 99.7% precision of 1 meter. After some research on the web you find out that you should use a Kalman filter to solve this task.

Define the components of your Kalman filter in the context of the state estimation of the pod. You can assume that the acceleration a is distributed normally around $\mu_a m/s^2$ with a variance equal to σ_a^2.

To define the Kalman filter we have to define the following 3 elements:

1. The prior (which is assumed to be Gaussian):
$$p(x_0) = \mathcal{N}(x_0|\mu_{x_0}, \Sigma_{x_0}).$$
2. The transition model (which is assumed to be Gaussian):
$$p(x_t|x_{t-1}) = \mathcal{N}(x_t|Ax_{t-1} + b, \Sigma_x).$$
3. The measurement model:
$$p(e_t|x_t) = \mathcal{N}(e_t|Cx_t + d, \Sigma_e).$$

In this case we can assume the initial position and speed of the pod are known with the same accuracy as the sensors gives, so we have the following definition for μ_{x_0} and Σ_{x_0}:

$$\mu_{x_0} = \begin{bmatrix} x_0 \\ \dot{x_0} \end{bmatrix}$$

and

$$\Sigma_{x_0} = \begin{bmatrix} (\frac{1}{3})^2 & 0 \\ 0 & (\frac{0.1}{3\sqrt{3}})^2 \end{bmatrix},$$

where x_0 and $\dot{x_0}$ respectively denote the initial position and speed of the pod. The covariance matrix values come from the fact that the 99.7% precision is equal to 3σ and that the variance of the mean of n random variables is equal to the sum of each variance divided by n^2.

Then to derive the transition model we use the laws of physics and we easily get:

$$A = \begin{bmatrix} 1 & \Delta_t \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} \frac{\mu_a}{2}\Delta_t^2 \\ \mu_a \Delta_t \end{bmatrix}$$

where Δ_t denotes the time elapsed between two measurements. The covariance matrix of the transition model can be computed by the affine transformation formula of a multivariate Gaussian which tells that if $y \sim \mathcal{N}(\mu_y, \Sigma_y)$ then $x = Qy + m \sim \mathcal{N}(\mu_x, \Sigma_x)$ where $\mu_x = Q\mu_y + m$ and $\Sigma_x = Q\Sigma_y Q^T$. Here we have $a \sim \mathcal{N}\mu_a, \sigma_a)$ which yields to :

$$\Sigma_x = \begin{bmatrix} \frac{1}{2}\Delta_t^2 \\ \Delta_t \end{bmatrix} \sigma_a^2 \begin{bmatrix} \frac{1}{2}\Delta_t^2 & \Delta_t \end{bmatrix} = \sigma_a^2 \begin{bmatrix} \frac{1}{4}\Delta_t^4 & \frac{1}{2}\Delta_t^3 \\ \frac{1}{2}\Delta_t^3 & \Delta_t^2 \end{bmatrix}.$$

Finally we can easily get that the mean of the sensors is given by:

$$C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad d = 0.$$

The covariance matrix is similarly computed and is equal to:

$$\Sigma_e = \begin{bmatrix} (\frac{1}{3})^2 & 0 & 0 & 0 \\ 0 & (\frac{0.1}{3})^2 & 0 & 0 \\ 0 & 0 & (\frac{0.1}{3})^2 & 0 \\ 0 & 0 & 0 & (\frac{0.1}{3})^2 \end{bmatrix}$$

[1] It means that 99.7% of the value measured will get a smaller error. e.g. (`https://math.stackexchange.com/questions/1412683/3-sigma-approximation`)