3

Gaussian Filters

3.1> Introduction

⇒ It is an important,
(family of recursive state estimation Algorithm).

⇒ Gaussian filters constitute the earliest tractable implementations of the Bayes filter for continuous spaces.

⇒ Gaussian techniques all share the basic idea that,

(beliefs are represented by multivariate normal distributions)

→ Gaussians are unimodal, {Single maximum}
- → Good for many tracking problem in robotics
- → Poor match for global estimation problem

⇒ The representation of Gaussian by its mean and Covariance is called the moments representation.

↳ Because mean and Covariance are the first and Second moments of a Probability distribution.

↳ All other moments are zero for normal distributions.

Canonical representation or Natural representation

↳ {alternate representation}

3.2) The Kalman Filter

3.2.1) Linear Gaussian Systems

⇒ Probably the best studied technique for implementing Bayes filters is the Kalman filter (KF).

⇒ Invented in 1950s by Rudolph Emil Kalman as a technique for filtering and prediction in linear systems.

⇒ It implements,
{belief computation for continuous states)

→ Not applicable to discrete or hybrid State Space

→ Posteriors are Gaussian if the following three properties hold in addition to the Markov assumptions of the Bayes filter:

1) The next state probability $p(x_t \mid u_t, x_{t-1})$ must be a linear function in its arguments with added Gaussian noise.

$$\boxed{x_t = A_t x_{t-1} + B_t u_t + \varepsilon_t}$$

→ Where, x_t and x_{t-1} are state vectors and u_t is control vector at time t.

$$x_t = \begin{pmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{n,t} \end{pmatrix} \quad \& \quad u_t = \begin{pmatrix} u_{1,t} \\ u_{2,t} \\ \vdots \\ u_{m,t} \end{pmatrix}$$

→ A_t and B_t are matrices.

→ The random variable ε_t is a Gaussian random vector that models the randomness in the state transition.

{ Its mean is zero and its covariance will be denoted by R_t }

Scanned by CamScanner

So,

$$\boxed{p(x_t \mid u_t, x_{t-1}) = \det(2\pi R_t)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x_t - A_t x_{t-1} - B_t u_t)^T R_t^{-1} (x_t - A_t x_{t-1} - B_t u_t)\right\}}$$

2) The measurement probability $p(z_t \mid x_t)$ must also be linear in its arguments, with added Gaussian noise.

$$\boxed{z_t = C_t x_t + \delta_t}$$ → describes measurement noise

→ C_t is a matrix of size $k \times n$, where k is the dimension of the measurement vector z_t.

→ δ_t has mean zero and covariance Q_t.

So,

$$\boxed{p(z_t \mid x_t) = \det(2\pi Q_t)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(z_t - C_t x_t)^T Q_t^{-1}(z_t - C_t x_t)\right\}}$$

3) Finally, the initial $bel(x_0)$ must be normal distributed. We will denote the mean of this belief by μ_0 and the covariance by Σ_0.

$$bel(x_0) = p(x_0) = \det(2\pi\Sigma_0)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x_0 - \mu_0)^T \Sigma_0^{-1}(x_0 - \mu_0)\right\}$$

3.2.2) The Kalman Filter Algorithm

$\Rightarrow$ Kalman filter represent the belief $bel(x_t)$ at time t by the mean μ_t and the Covariance Σ_t.

Algorithm Kalman-filter $(\mu_{t-1}, \Sigma_{t-1}, u_t, z_t)$:

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

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

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

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

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

return μ_t, Σ_t

{Kalman gain}

$\Rightarrow$ The Kalman filter is computationally quite efficient.

3.3) The Extended Kalman filter (EKF)

⇒ The assumptions of linear state transitions and linear measurements with added Gaussian noise are rarely fulfilled in Practice.

⇒ The extended Kalman filter (EKF) overcomes this assumption.

⇒ Here the assumption is that the next state Probability and the measurement probabilities are governed by nonlinear functions g and h, respectively.

$$x_t = g(u_t, x_{t-1}) + \epsilon_t$$

$$z_t = h(x_t) + \delta_t$$

⇒ Unfortunately, with arbitrary functions g and h, the belief is no longer a Gaussian.

↳ In fact, performing the belief update exactly is usually impossible for nonlinear functions g and h, in the sense that the Bayes filter does not possess a closed-form Solution.

⇒ The EKF calculates an approximation to the true belief.

↳ It represents this approximation by a gaussian.

3.3.D Linearization via Taylor Expansion

⇒ The Key idea underlying the EKF is called Linearization.

⇒ There exist many techniques for linearizing nonlinear functions.

↳ EKFs utilize a method called (first order) Taylor expansion. (about mean)

$$g(u_t, x_{t-1}) \approx g(u_t, \mu_{t-1}) + g'(u_t, \mu_{t-1})(x_{t-1} - \mu_{t-1})$$

$$= g(u_t, \mu_{t-1}) + G_t(x_{t-1} - \mu_{t-1})$$

⇒ The next state probability is approximated as follows:

$$\boxed{p(x_t | u_t, x_{t-1}) \approx \det(2\pi R_t)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}[x_t - g(u_t, \mu_{t-1}) - G_t(x_{t-1} - \mu_{t-1})]^T R_t^{-1} [x_t - g(u_t, \mu_{t-1}) - G_t(x_{t-1} - \mu_{t-1})]\right\}}$$

⇒ G_t is a matrix of size $n \times n$.

→ also called Jacobian.

$$\Rightarrow h(x_t) \approx h(\bar{\mu}_t) + h'(\bar{\mu}_t)(x_t - \bar{\mu}_t)$$
$$= h(\bar{\mu}_t) + H_t(x_t - \bar{\mu}_t)$$

$\Rightarrow$ Written in a Gaussian, we have,

$$p(z_t | x_t) = \det(2\pi Q_t)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}[z_t - h(\bar{\mu}_t) - H_t(x_t - \bar{\mu}_t)]^T Q_t^{-1}[z_t - h(\bar{\mu}_t) - H_t(x_t - \bar{\mu}_t)]\right\}$$

3.3.2) The EKF Algorithm

Algorithm Extended Kalman filter $(\mu_{t-1}, \Sigma_{t-1}, u_t, z_t)$:

$$\bar{\mu}_t = g(u_t, \mu_{t-1})$$
$$\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T + R_t$$
$$K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + Q_t)^{-1}$$
$$\mu_t = \bar{\mu}_t + K_t(z_t - h(\bar{\mu}_t))$$
$$\Sigma_t = (I - K_t H_t)\bar{\Sigma}_t$$

return μ_t, Σ_t

$\Rightarrow$ EKF has become just about the most popular tool for state estimation in robotics

$\Rightarrow$ Its strength lies in its simplicity and in its computational efficiency.

⇒ EKFs are incapable of representing multimodal beliefs.

↳ A common extension of EKFs is to represent posteriors using mixtures or sum of Gaussians

↳ EKFs that utilize such mixture representations are called <u>multi-hypothesis (extended) Kalman filter</u> {MHEKF}

3.4) <u>The Information Filter</u>

⇒ The dual of Kalman filter is the information filter.

↳ IF also represents the belief by a Gaussian

⇒ The key difference between KF and the IF arises from the way the Gaussian belief is represented.

↳ Information filters represent Gaussians in their canonical representation, which is comprised of

<u>Information matrix</u> <u>Information Vector</u>

⇒ The difference in representation leads to different update equations.

↳ which makes it computationally simple.

3.4.1) Canonical Representation

⇒ The canonical representation of a multivariate Gaussian is given by a matrix Ω and a vector ξ.

$$\Omega = \Sigma^{-1} \quad \{\text{Information matrix or Precision matrix}\}$$

$$\xi = \Sigma^{-1}\mu \quad \{\text{Information Vector}\}$$

$$\Rightarrow P(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$

$$= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu - \frac{1}{2}\mu^T\Sigma^{-1}\mu\right\}$$

$$= \underbrace{\det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\mu^T\Sigma^{-1}\mu\right\}}_{\text{Const}} \exp\left\{-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu\right\}$$

$$\Rightarrow P(x) = \eta \exp\left\{-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu\right\}$$

$$\Rightarrow \boxed{P(x) = \eta \exp\left\{-\frac{1}{2}x^T\Omega x + x^T\xi\right\}}$$

$$-\log P(x) = \text{const} + \frac{1}{2}x^T\Omega x - x^T\xi$$

3.4.2) The Information Filter Algorithm

Algorithm Information_filter $(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)$:

$$\bar{\Omega}_t = (A_t \Omega_{t-1}^{-1} A_t^T + R_t)^{-1}$$

$$\bar{\xi}_t = \bar{\Omega}_t (A_t \Omega_{t-1}^{-1} \xi_{t-1} + B_t u_t)$$

$$\Omega_t = C_t^T Q_t^{-1} C_t + \bar{\Omega}_t$$

$$\xi_t = C_t^T Q_t^{-1} z_t + \bar{\xi}_t$$

return ξ_t, Ω_t

3.4.4) The Extended Information Filter Algorithm

Algorithm Extended_information_filter $(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)$:

$$\mu_{t-1} = \Omega_{t-1}^{-1} \xi_{t-1}$$

$$\bar{\Omega} = (G_t \Omega_{t-1}^{-1} G_t^T + R_t)^{-1}$$

$$\bar{\xi}_t = \bar{\Omega}_t \, g(u_t, \mu_{t-1})$$

$$\bar{\mu}_t = g(u_t, \mu_{t-1})$$

$$\Omega_t = \bar{\Omega}_t + H_t^T Q_t^{-1} H_t$$

$$\xi_t = \bar{\xi}_t + H_t^T Q_t^{-1} [z_t - h(\bar{\mu}_t) - H_t \bar{\mu}_t]$$

return ξ_t, Ω_t