Statistical Techniques in Robotics (16-831, F10) Lecture #19 (Tuesday October 26)

Kalman Filtering (part 1)

Lecturer: Drew Bagnell *Scribe: Kiho Kwak, Siddharth Mehrotra*

Gauss Markov Filter

Consider $X_1, X_2, ...X_t, X_{t+1}$ to be the state variables and $Y_1, Y_2, ...Y_t, Y_{t+1}$ be the sequence of corresponding observations. As in Hidden Markov models, conditional independencies (see Figure 1) dictate that past and future states are decorrelated given the current state, X_t at time t. For example, if we know what X_2 is, then no information about X_1 can possibly help us to reason about what X_3 should be.

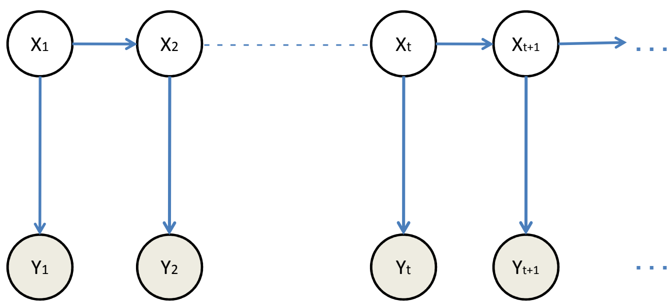

Figure 1: The Independence Diagram of a Gauss-Markov model

The update for state variable X_{t+1} is given by:

$$X_{t+1} = AX_t + \epsilon$$

where,

$$\epsilon \sim N(0, Q)$$

$$\Rightarrow X_{t+1}|X_t \sim N(AX_t, Q)$$

The corresponding observation Y_{t+1} is given by equation:

$$Y_{t+1} = CX_{t+1} + \delta$$

where,

$$\delta \sim N(0, R)$$

$$\Rightarrow Y_0 \sim N(\mu_0, \epsilon_0)$$

Each component is defined as follow:

- A_t: Matrix (nxn) that describes how the state evolves from t to t-1 without controls or noise.
- C_t: Matrix (kxn) that describes how to map the state X_t to an observation Y_t.

- ϵ_t, δ_t: Matrix (nxn) Random variables representing the process and measurement noise that are assumed to be independent and normally distributed with covariance R_t and Q_t respectively.

Lazy Gauss Markov Filter

Motion Model (Prediction step):

Before the observation is taken:

$$X_{t+1} \sim \mu_{t+1}^{-} = A\mu_t$$

Proof:

$$E[X_{t+1}] = E[AX_t + \epsilon]$$
$$\Rightarrow E[X_{t+1}] = E[AX_t] + E[\epsilon]$$

since variance of ϵ is 0,

$$\Rightarrow E[X_{t+1}] = AE[X_t] = A\mu_t$$

Variance,

$$\Sigma_{t+1}^{-} = E[X_{t+1} * {X_{t+1}}^T]$$
$$\Rightarrow \Sigma_{t+1}^{-} = E[(AX_t + \epsilon)(AX_t + \epsilon)^T]$$
$$= E[(AX_t)(AX_t)^T] + E[\epsilon_{terms}]$$
$$= AE[(X_t)(X_t)^T]A^T + E[\epsilon_{terms}]$$
$$\Rightarrow \Sigma_{t+1}^{-} = A\Sigma_t A^T + E[\epsilon_{terms}]$$

E[ϵ_{terms}] is equal to the variance of ϵ which is Q.
Therefore Motion Update becomes:

$$\mu_{t+1}^{-} = A\mu_t$$
$$\Sigma_{t+1}^{-} = A\Sigma_t A^T + Q$$

Observation Model (Correction step):

For the observation model Natural parameterization is more suitable as it involves multiplication of terms. When, Y is the corresponding observation for state variable X, the model equation in terms of Natural Parameters J and P is given by,

$$e^{(J^{-T}X - \frac{1}{2}X^T PX)} * e^{-\frac{1}{2}(Y-CX)^T R^{-1}(Y-CX)}$$

$$\Rightarrow e^{-\frac{1}{2}[-2Y^T R^{-1}CX + X^T C^T R^{-1}CX + Y^T R^{-1}Y]}$$

The last term is a constant with respect to X, so it goes into the marginalization term.

$$\Rightarrow e^{-\frac{1}{2}[-2Y^TR^{-1}CX+X^TC^TR^{-1}CX]}$$

Therefore the Observation Update is:

$$J^+ = J^- + (Y^TR^{-1}C)^T$$
$$P^+ = P^- + C^{-1}R^{-1}C$$

This form is useful when there are large number of motion and observation updates.
Lazy Gauss Markov can be expressed in two forms:

- When expressed in terms of moment parameters μ and Σ acts as **Kalman Filter**.
- When expressed in terms of natural parameters J and P acts as **Information Filter**.

Observation Update in terms of moment parameters μ and Σ:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{X_t} \\ \mu_{Y_t} \end{pmatrix} \begin{pmatrix} \Sigma_{XX} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{YY} \end{pmatrix} \right)$$

Observation Update:

$$\mu_{X|Y} = \mu_X + \Sigma_{XY}\Sigma_{YY}^{-1}(Y - \mu_Y) \leftarrow (1)$$
$$\Sigma_{X|Y} = \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX} \leftarrow (2)$$

$\Sigma_{XY}\Sigma_{YY}^{-1}$ is Kalman gain, K_t and $(Y - \mu_Y)$ is called the Innovation Term.

We know,

$$\mu_Y = C\mu_X$$
$$\Sigma_{YY} = R + C\Sigma_{XX}C^T$$

Therefore, we have to find out Σ_{XY} to calculate the remaining terms in equation 1 and 2.
By definition,

$$\Sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)^T]$$
$$= E[(X - \mu_X)(Y - C\mu_X)^T]$$

However, $Y = CX + \delta$ with δ having 0 mean and independent of all other observations.

$$\Sigma_{XY} = E[(X - \mu_X)(X - \mu_X)^T]C^T$$

$$\Rightarrow \Sigma_{XY} = \Sigma_{XX}C^T$$

Putting, these values in equations 1 and 2,

$$\mu_{X|Y} = \mu_X + \Sigma_{XX}C^T(R + C\Sigma_{XX}C^T)^{-1}(Y - C\mu_X)$$

$$\Sigma_{X|Y} = \Sigma_{XX} - \Sigma_{XX}C^T(R + C\Sigma_{XX}C^T)^{-1}C\Sigma_{XX}$$

Algorithm **Kalman_filter**(μ_{t-1}, Σ_{t-1}, u_t, z_t):

Prediction:

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

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

Correction:

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

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

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

Return μ_t, Σ_t

Figure 2: The Kalman filter algorithm

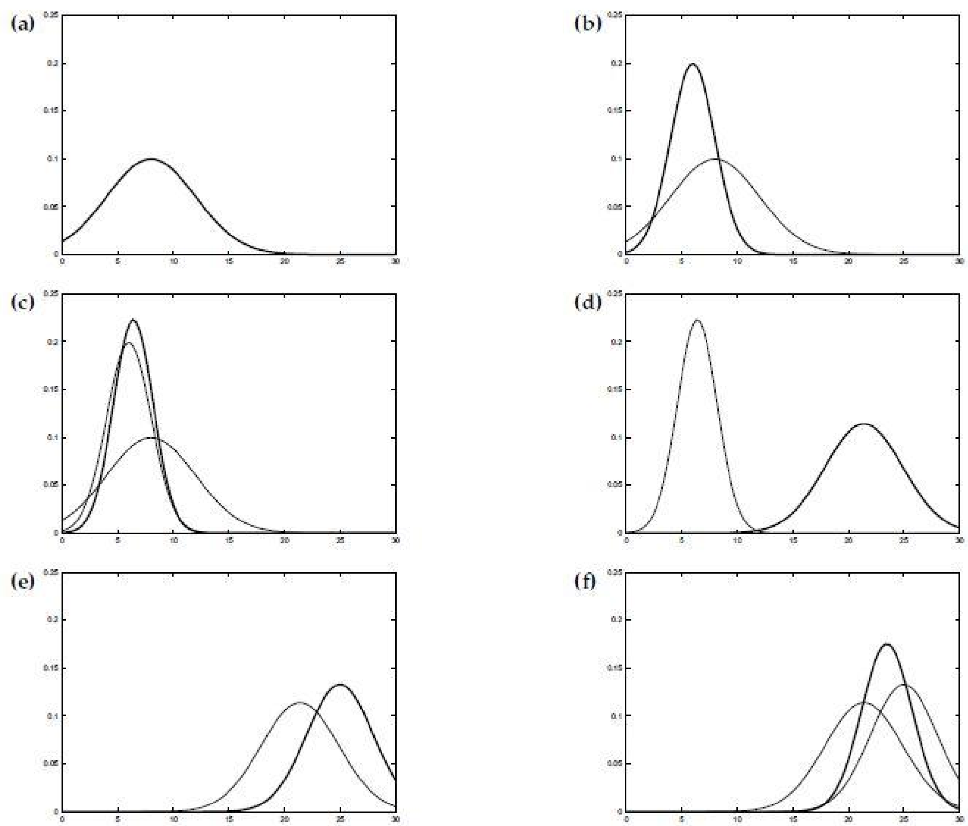

Figure 3: Illustration of Kalman filters: (a) initial belief, (b) a measurement (in bold) with the associated uncertainty, (c) belief after integrating the measurement into the belief using the Kalman filter algorithm, (d) belief after motion to the right (which introduces uncertainty), (e) a new measurement with associated uncertainty, and (f) the resulting belief.